AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

R. BRAUER
UNIVERSITY OF TORONTO

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

L. M. GRAVES
UNIVERSITY OF CHICAGO

H. WHITNEY
HARVARD UNIVERSITY

A. WINTNER
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

S. BOCHNER	G. BIRKHOFF	L. BERS
D. MONTGOMERY	R. COURANT	J. L. DOOB
C. L. SIEGEL	P. HARTMAN	N. JACOBSON
D. C. SPENCER	J. L. VANDERSLICE	D. C. LEWIS
M. A. ZORN	A. ZYGMUND	A. D. WALLACE

PUBLISHED UNDER THE JOINT AUSPICES OF
THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LXIX

1947

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

Mathematics Library
D
MAR 1 8 1947

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

R. BRAUER
UNIVERSITY OF TORONTO

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

L. M. GRAVES
UNIVERSITY OF CHICAGO

H. WHITNEY
HARVARD UNIVERSITY

A. WINTNER
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

S. BOCHNER	G. BIRKHOFF	L. BERS
D. MONTGOMERY	R. COURANT	J. L. DOOB
C. L. SIEGEL	P. HARTMAN	N. JACOBSON
D. C. SPENCER	J. L. VANDERSLICE	D. C. LEWIS
M. A. ZORN	A. ZYGMUND	A. D. WALLACE

PUBLISHED UNDER THE JOINT AUSPICES OF
THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXIX, Number 1
JANUARY, 1947

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

CONTENTS

	PAGE
The free product of groups. By EMIL ARTIN,	1
(L^2)-connections between the potential and kinetic energies of linear systems. By AUREL WINTNER,	5
Identities in the theory of power series. By ISSAI SCHUR,	14
A note on division rings. By N. JACOBSON,	27
The double chain condition in cyclic operator groups. By REINHOLD BAER,	37
Radicals and subdirect sums. By BAILEY BROWN and NEAL H. MCCOY,	46
On congruences and conjugate nets. By V. G. GROVE,	59
The elliptic modular function and a class of analytic functions first considered by Hurwitz. By ZEEV NEHARI,	70
On the Laplace-Fourier transcendents occurring in mathematical physics. By AUREL WINTNER,	87
On the Tauberian nature of Ikehara's theorem. By AUREL WINTNER,	99
Union-preserving transformations of higher order surface-elements. By JOHN DE CICCO,	104
A new definition of the Godeaux sequence of quadrics. By CHENKUO PA,	117
Unilateral variations with variable end-points. By JULIAN D. MANCILL,	121
On some triangular summability methods. By JOSHUA BARLAZ,	139
Topological rings. By IRVING KAPLANSKY,	153
Uniqueness of solutions of ultrahyperbolic partial differential equations. By O. G. OWENS,	184
Note on some strong laws of large numbers. By KAI-LAI CHUNG,	189

The AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications may be sent to Professor F. D. MURNAGHAN at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE 18, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

THE FREE PRODUCT OF GROUPS.*

By EMIL ARTIN.

The traditional way of introducing the free product of groups consists in defining carefully the product of elements and proving the associative law. We shall use here, on the contrary, a method where the associative law is obvious and all the difficulties arise from the discussion of equality.

Let $\mathfrak{S}$ be a set of groups Γ. Without loss of generality, we assume these groups to be disjoint. The elements of all these groups will be denoted by small letters $a, b, c, \cdots$ whereas the capitals $A, B, \cdots$ stand for symbolic products

$$A = a_1 a_2 \cdots a_n. \tag{1}$$

The elements a_i may or may not belong to the same Γ. The number n of terms is called the length of A. Two such products are equal if and only if they consist of the same factors in precisely the same order; in other words, no simplification is allowed in (1), not even if two adjoining elements belong to the same Γ or if one of the factors is a unit. Multiplication AB is defined in the obvious way by writing first the factors of A and then those of B. The associative law is therefore trivial. It is convenient to introduce also the "empty" product E which acts as unit element. Nevertheless we do not get a group since $A \neq E$ does not have an inverse element. In spite of this we use as pure notation A^{-1} for the product $a_n^{-1} a_{n-1}^{-1} \cdots a_1^{-1}$.

It might happen that all the a_i in (1) are from the same group Γ and that the product $a_1 a_2 \cdots a_n$ computed in Γ has the value 1. We then call A an elementary expression.

Let us now define a certain set Δ of formal products. It contains those products that we want later to equate to E and is defined by induction:

1) E is in Δ.

2) If A has a positive length it belongs to Δ if and only if it can be written in the form

$$A = A_0 a_1 A_1 a_2 \cdots A_{n-1} a_n A_n \tag{2}$$

* Received May 10, 1946.

where all A_i are elements of Δ (and have smaller length, of course) and where the a_i are such that $a_1 a_2 \cdots a_n$ is an elementary expression.

We prove now a few lemmas about our set Δ.

1) $A \in \Delta$ implies $A^{-1} \in \Delta$; $A \in \Delta$ and $B \in \Delta$ implies $AB \in \Delta$. Proof by induction. Both parts are true for $A = E$. Let A be of the form (2). Every A_i is shorter than A so that the theorem may be assumed for A_i. This proves directly the first half. For the second half we can assume $A_n B \in \Delta$. Now AB has again the form (2) with $A_n B$ instead of A_n.

2) If $A \in \Delta$ and $A \neq E$ we can always assume that A is of the form (2) with $A_0 = E$ (or also with $A_n = E$ if we prefer). Proof by induction. Assume A of form (2) and denote by B the whole product (2) but without A_0. Then $B \in \Delta$ and $A = A_0 B$. If $A_0 = E$ we have nothing to prove; at any rate A_0 is shorter than A so that we may assume $A_0 = b_1 B_1 b_2 B_2 \cdots B_n$. Now $A = b_1 B_1 b_2 B_2 \cdots b_n (B_n B)$.

3) If the first and the last term of an $A \in \Delta$ belong to the same Γ, we can even assume in (2) that $A_0 = A_n = E$.

Proof. We can assume $A_0 = E$; now $A_n \in \Delta$ so that we can write A_n in the form

$$A_n = B_0 b_1 B_1 b_2 \cdots b_m.$$

This leads to

$$A = a_1 A_1 a_2 A_2 \cdots a_n B_0 b_1 B_1 b_2 \cdots B_{m-1} b_m.$$

Since $a_1 a_2 \cdots a_n b_1 \cdots b_m$ is elementary, A is of the required form.

4) aAa^{-1} and A belong to Δ or do not at the same time. *Proof.* a) Let $A \in \Delta$. aa^{-1} is elementary so $aAa^{-1} \in \Delta$. b) Let $aAa^{-1} \in \Delta$. According to a previous lemma it is of the form

$$aAa^{-1} = aA_1 a_2 A_2 \cdots a_{n-1} A_{n-1} a^{-1}$$

so that we have

$A = A_1 a_2 A_2 \cdots a_{n-1} A_{n-1}$ and $a_2 a_3 \cdots a_{n-1}$ is elementary.

Repeated application proves:

5) BAB^{-1} and A belong to Δ or do not at the same time.

6) $AB \varepsilon \Delta$ implies $BA \varepsilon \Delta$. *Proof.* $AB \varepsilon \Delta$ implies $AB \cdot AA^{-1} \varepsilon \Delta$ since $AA^{-1} = AEA^{-1}$ belongs to Δ. From 5) we get now $BA \varepsilon \Delta$.

DEFINITION. A and B are called congruent, $A \equiv B$, if and only if $AB^{-1} \varepsilon \Delta$ ($B^{-1}A \varepsilon \Delta$ would mean the same).

We see at once:

a) $A \equiv A$ since $AA^{-1} \varepsilon \Delta$.

b) $A \equiv B$ implies $B \equiv A$. Indeed $AB^{-1} \varepsilon \Delta$ implies $BA^{-1} \varepsilon \Delta$.

c) $A \equiv B$ and $B \equiv C$ implies $A \equiv C$, because $AB^{-1} \varepsilon \Delta$ and $BC^{-1} \varepsilon \Delta$ imply $BC^{-1}AB^{-1} \varepsilon \Delta$ and therefore $C^{-1}A \varepsilon \Delta$.

d) $A \equiv B$ and $C \equiv D$ imply $AC \equiv BD$. Indeed $B^{-1}A \varepsilon \Delta$ and $CD^{-1} \varepsilon \Delta$ imply $B^{-1}ACD^{-1} \varepsilon \Delta$ and therefore $ACD^{-1}B^{-1} \varepsilon \Delta$.

e) $AA^{-1} \equiv E$.

THEOREM 1. *Our formal products form a group if we use the congruence instead of equality. The elements of Δ are precisely those congruent to E.*

Let us now consider the elements $A = a$ of length 1. Two of them are congruent, $a \equiv b$, if and only if ab^{-1} is in Δ. Either ab^{-1} has to be elementary, which means $a = b$, or both factors have to be in Δ and therefore to be the unit elements of their group. Our congruence equates, therefore, only the unit elements among our elements a. We therefore denote from now on all unit elements by 1. Let next a and b belong to the same group Γ and let c be their product in Γ. The formal product ab is then congruent to c because abc^{-1} is elementary.

Our group contains, therefore, each of the groups Γ as a subgroup, and different Γ's have only the unit in common.

In the sense of congruence we are now in the position to simplify a symbolic product A by uniting two adjoining factors as soon as they belong to the same group and by dropping every unit element. Since these operations reduce the length of A we finally arrive either at E or at an expression where no factor is 1 and no two adjoining terms are of the same group. We call such an expression normal.

THEOREM 2. *Every A is congruent to a normal form.*

LEMMA. *A normal form $A = a_1 a_2 \cdots a_n$ of length > 0 cannot be $\equiv 1$.*

Proof by induction. Assume it true for shorter expressions. If we write A in the form (2) all the A_i are shorter than A and also normal (as parts of A). So they must be E and only $A = a_1 a_2 \cdots a_n$ remains and has to be elementary. But that is impossible.

If we ask now when two normal expression can be congruent, $a_1 a_2 \cdots a_n \equiv b_1 b_2 \cdots b_m$, we have to consider whether $a_1 a_2 \cdots a_n b_m^{-1} b_{m-1}^{-1} \cdots b_1^{-1}$ can be $\equiv 1$. Since this product must never become a normal expression before being reduced to E we easily deduce:

THEOREM 3. *Two normal expressions are congruent if and only if they are identical.*

PRINCETON UNIVERSITY,
PRINCETON, N. J.

(L^2)-CONNECTIONS BETWEEN THE POTENTIAL AND KINETIC ENERGIES OF LINEAR SYSTEMS.*

By AUREL WINTNER.

1. Let primes denote differentiations with respect to a real variable, t, the range of which will be the half-line $t \geqq 0$. The functions to be considered, $f(t), g(t), x(t)$, will be assumed to be defined and continuous on this half-line. For the sake of simplicity, it will be assumed from the beginning that all these functions are scalars. In addition, *all* functions will be regarded as real-valued.

Correspondingly, $x(t)$ will be said to be of class (L^2) if

$$\int_0^\infty x^2(t)\,dt < \infty.$$

This is neither necessary nor sufficient in order that $x(t)$ be of class $(L) = (L^1)$:

$$\int_0^\infty |x(t)|\,dt < \infty.$$

In fact, one of the principal difficulties connected with the oscillation problems to be considered is that

$$x(t) \to 0 \text{ as } t \to \infty \tag{1}$$

cannot be assumed.

The results will deal with the (L^2)-solutions, $x(t)$, of linear differential equations

$$x'' + f(t)x = 0 \tag{2}$$

and of the corresponding inhomogeneous equations

$$x'' + f(t)x = g(t). \tag{3}$$

The issue will be the (L^2)-character of the derivative, $x'(t)$. Such problems suggest themselves by the usual formal treatment of the characteristic functions, $x(t)$, of a Schrödinger equation (where $t = r$). For, on the one hand, neither

* Received September 25, 1946.

$$\int_0^\infty x'^2(t)\,dt < \infty$$

nor

$$x'(t) \to 0 \text{ as } t \to \infty$$

should be assumed of a characteristic function and, on the other hand, the partial integrations (on which, for instance, the deduction of the relation of indeterminacy depends) cannot be made legitimate otherwise.

In order to avoid this difficulty, I gave (iii) below, under the restriction (4), in a lecture many years ago. I noticed only recently that what is actually needed when "locating" continuous spectra, namely, (iv) below, can be concluded from (iii).

2. If (2) is interpreted as a dynamical system, then $\frac{1}{2}x'^2$ is the kinetic, and $\frac{1}{2}fx^2$ the potential, energy of a solution $x = x(t)$. Their sum is not a constant, since the Lagrangian equation, (2), contains t explicitly. To say that $x'(t)$ is of class (L^2) means that the sum total of the *density* of the kinetic energy remains finite as $t \to \infty$. The corresponding remark for $x(t)$ and the potential energy, $\frac{1}{2}fx^2$, holds only if $|f|^{\frac{1}{2}}$, the absolute frequency of (2), does not come "too close" to 0 and ∞, as $t \to \infty$.

This suggests that, under an appropriate restriction of the behavior of $f(t)$ as $t \to \infty$, the (L^2)-character of $x'(t)$ can be deduced from that of $x(t)$, if use is made of a form of those "virial" considerations which, in the latter part of the fourth of his *Vorlesungen über Dynamik,* Jacobi has applied to the (non-linear, but conservative) case of a solar system. What can actually be proved in this manner is the second of the assertions of the following theorem:

(i) *If, as* $t \to \infty$,

$$f(t) = O(1), \tag{4}$$

then a solution, $x(t)$, *of* (2) *cannot be of class* (L^2) *unless* $x'(t)$ *is also of class* (L^2). *The same is true if* (4) *is relaxed to*

$$f(t) = O_R(1). \tag{5}$$

It is not in general true if (4) *is omitted entirely; in fact, even* $f(t) > 0$, *hence*

$$f(t) = O_L(1), \tag{5 bis}$$

is insufficient.

In (5) and (5 bis), the subscripts refer to boundedness from above ("right") and from below ("left"), respectively.

The insufficiency of the restriction (5 bis) is proved by the example

$$(5^*) \qquad f(t) = (e^t + \tfrac{1}{2})(e^t - \tfrac{1}{2}) > 0, \qquad (0 \leqq t < \infty).$$

For, in this case, (2) is readily verified to have the solution $x(t) = e^{-\frac{1}{2}t} \cos e^t$. Clearly, this solution is of class (L^2), and (5 bis) is satisfied. Nevertheless, $x'(t)$ is not of class (L^2). In fact, $x'(t)$ is the sum of two terms, one of which is $O(e^{-\frac{1}{2}t})$, hence a function of class (L^2), whereas the other term is $-e^{\frac{1}{2}t} \sin e^t$, a function which, in view of

$$\int_0^\infty (-e^{\frac{1}{2}t} \sin e^t)^2 dt = \int_1^\infty \sin^2 s \, ds = \infty, \qquad (s = e^t),$$

fails to be of class (L^2).

3. The verification of the preceding solution makes it clear that, if $f(t)$ is given by (5*), both

$$x(t) = e^{-\frac{1}{2}t} \cos e^t \quad \text{and} \quad x(t) = e^{-\frac{1}{2}t} \sin e^t$$

are solutions of (2). Since both of these solutions are of class (L^2), the example (5*) proves the second of the assertions of the following theorem:

(ii) *If $f(t)$ is non-negative and satisfies, uniformly for all large t, a Lipschitz condition,*

$$|f(t_1) - f(t_2)| \leqq \text{const.} \, |t_1 - t_2|,$$

then (2) cannot have two linearly independent solutions of class (L^2). The same is not in general true if the Lipschitz condition (which, incidentally, allows unbounded functions $f(t)$, such as $t^{\frac{1}{2}}$ or t, and even $t \sin^2 1/t$) is omitted, that is, if only (5 bis) or $f(t) \geqq 0$ is assumed.

On the other hand, the Lipschitz condition is superfluous if (5 bis) is replaced by the assumption, (5), of (i). More generally, (2) cannot have two linearly independent solutions of class (L^2) if $f(t)$ has the property that no solution, $x(t)$, of (2) is of class (L^2) unless $x'(t)$ also is of class (L^2).

The proof of the first assertion of (ii) proceeds as follows:

It is easily verified (for instance, by a partial integration of the relevant Stieltjes integral) that

$$d\{x'^2(t) + f(t)x^2(t)\} = x^2(t) df(t)$$

is an identity in t along any solution, $x(t)$, of (2) (the notation $df(t)$ is justified in itself, since $f(t)$, being subject to a Lipschitz condition, is absolutely continuous). Clearly, the last two formula lines imply the inequality

$$\int^{\infty} |\, d\{x'^2(t) + f(t)x^2(t)\}\,| \leqq \text{const.} \int^{\infty} x^2(t)\,dt.$$

Hence, if $x(t)$ is of class (L^2), the integral on the left of this inequality is convergent. This means that the function $x'^2(t) + f(t)x^2(t)$ is of bounded variation. In particular, it is a bounded function. It follows therefore from the first assumption, $f(t) \geqq 0$ (which has not been used thus far), that $x'^2(t)$ is bounded. In other words, a solution, $x(t)$, cannot be of class (L^2) unless $x'(t) = O(1)$.

Consequently, if $x = x(t)$ and $x = y(t)$ are two solutions of class (L^2), then $x'(t) = O(1)$ and $y'(t) = O(1)$, and so the Wronskian, $x'(t)y(t) - y'(t)x(t)$, is of class (L^2). But the Wronskian of any two solutions of (2) is independent of t (Abel). Since a function which is independent of t is of class (L^2) on the half-line $0 \leqq t < \infty$ only if it vanishes identically, it follows that the Wronskian of $x(t)$ and $y(t)$ is 0. Accordingly, $x(t)$ and $y(t)$ are linearly dependent. This proves the first assertion of (ii).

The fourth assertion of (ii) (which, in view of (i), implies the third), can be concluded by a slight modification of the last argument. For, if $x = x(t)$ and $y = y(t)$ are of class (L^2), then, by the present assumption, the same is true of $x'(t)$ and $y'(t)$. But the product of two functions of class (L^2) is absolutely integrable (Schwarz), hence integrable. Consequently, the Wronskian, although a constant, is integrable on the half-line $0 \leqq t < \infty$ and must, therefore, vanish. Since this means that $x(t)$ and $y(t)$ are linearly dependent, the last assertion of (ii) follows.

4. The third assertion of (i) has been proved by the example (5*). The second assertion of (i), which remains to be proved, implies the following

COROLLARY. *If $f(t)$ is restricted by* (5), *then a solution, $x(t)$, of* (2) *cannot be of class (L^2) unless $x(t)$ satisfies* (1).

First, (i) and the assumptions of the Corollary imply that both functions $x(t)$, $x'(t)$ are of class (L^2). Hence, the product $x(t)x'(t)$ is absolutely integrable, and so its integral over the half-line $0 \leqq t < \infty$ is convergent. But, since $xx' = (\frac{1}{2}x^2)'$, the convergence of this improper integral means that

the square of $x(t)$ tends to a finite limit as $t \to \infty$. Finally, this limit cannot be distinct from 0, since $x(t)$ is of class (L^2).

5. The remaining part of (i) is a particular case, $g \equiv 0$, of the following theorem:

(iii) *If $f(t)$ is restricted by* (4), *or merely by* (5), *and if $g(t)$ is of class (L^2), then a solution, $x(t)$, of* (3) *cannot be of class (L^2) unless $x'(t)$ is also of class (L^2).*

Only the "external force," $g(t)$, and not the main coefficient function, $f(t)$, of (2) as well, is required to be of class (L^2).

Since x and g are supposed to be of class (L^2), the product xg is of class (L). Hence, (3) shows that the sum $xx'' + fx^2$ is of class (L). In order to avoid a secondary complication which arises when (4) is generalized to (5), suppose first (4). Then, since x is of class (L^2), hence x^2 of class (L), the product fx^2 is of class (L). But the same is true of the sum $xx'' + fx^2$. Consequently, xx'' is of class (L).

However, xx'' is identical with $(xx')' - x'^2$. On the other hand, the assertion of (iii), according to which x' is of class (L^2), means that x'^2 is of class (L). Hence, if the assertion were not true, then, since $(xx')' - x'^2$ is of class (L), and since $x'^2 \geqq 0$, it would follow that the indefinite integral of $(xx')'$ must *tend* to ∞ as $t \to \infty$. But this leads to a contradiction.

In fact, the indefinite integral just mentioned is $xx' + \text{const}$. Since $xx' = (\frac{1}{2}x^2)'$, it follows that what is supposed to *tend* to ∞ is $(\frac{1}{2}x^2)'$. But a quadrature shows that $(x^2)' \to \infty$ implies $x^2 \to \infty$, as $t \to \infty$. However, $x^2 \to \infty$ contradicts the assumption that $x(t)$ is of class (L^2). This proves (iii) in the particular case, (4), of (5).

In order to relax (4), put $f(t) = p(t) + n(t)$, where $p = \max(f, 0)$. Then both functions p, $-n$ are non-negative (and continuous, since f is). Furthermore, since the assumption (5) means that p is bounded, and since x is of class (L^2), the product px^2 is of class (L). But $x'' + fx^2$ was seen to be of class (L) (without any assumption on f). It follows therefore from $fx^2 = px^2 + nx^2$ that the sum $xx'' + nx^2$ is of class (L).

In particular, the indefinite integral of this sum tends to a finite limit as $t \to \infty$. If the indefinite integral of the second term, nx^2, of the sum does not tend to a finite limit, then, since $nx^2 \leqq 0$, it must *tend* to $-\infty$, and so the indefinite integral of the first term, xx'', must *tend* to ∞. Since xx'', being identical with $(xx')' - x'^2$, is not less than $(xx')'$, this leads to the same contradiction as before.

Consequently, the indefinite integral of nx^2 cannot tend to $-\infty$. Since $nx^2 \leqq 0$, and since $xx'' + nx^2$ is of class (L), it follows that xx'' is of class (L). The balance of the proof is the same as before.

6. It will be convenient to use the following abbreviated manner of speaking:

DEFINITION. *Let $f(t)$ be called of class (*) if every solution of the corresponding homogeneous differential equation, (2), satisfies*

$$x(t) = O(1) \text{ and } x'(t) = O(1) \qquad (t \to \infty). \tag{6}$$

There are various explicit criteria sufficient in order that $f(t)$ be of class (*). For instance, it is known that $f(t)$ is sure to be of class (*) if there exists a positive constant, ω^2, satisfying either

$$\int^{\infty} | f(t) - \omega^2 | \, dt < \infty$$

(where $0 < \limsup_{t\to\infty} | f(t) - \omega^2 | \leqq \infty$ is allowed) or

$$\int^{\infty} | df(t) | < \infty \text{ and } f(t) \to \omega^2 \text{ as } t \to \infty$$

(these two criteria are independent, even if $f(t) \to \omega^2$ is assumed in the first). However, in the following theorem, the property (*) of $f(t)$ can be thought of as assured by *some* criterion.

(iv) *Let $f(t)$ be a function which is of class (*) and satisfies (4) or, more generally, (5), and let $g(t)$ be a* fixed *function of class (L^2). Then the inhomogeneous equation (3), belonging to the given $g(t)$, possesses a solution of class (L^2) if and only if there belongs to every solution, $x(t)$, of the homogenous equation, (2), a (unique) constant, $c = c_{x(t)}$, for which the function*

$$c - \int_0^t x(u)g(u)du \tag{8}$$

becomes a function of class (L^2).

The existence of such a constant c is required for every solution of (2), rather than for only those solutions of (2) which are of class (L^2).

The point in (iv) is that, in contrast to what is usually assumed in the theory of spectra, the (L^2)-requirement is now placed only on a solution of

the homogeneous equation, rather than on such a solution *and* on its derivative. The rôle of (iii) is precisely this reduction.

7. The proof of (iv) proceeds as follows:

Let $x(t)$ denote a *fixed* solution of (2). If $x(t) \equiv 0$, then the constant the existence of which is claimed in (iv) clearly is $c = 0$. It can therefore be assumed that $x(t) \not\equiv 0$. Then (2) has a solution, say $y(t)$, which is linearly independent of $x(t)$. This means that the determinant of the Wronksian matrix,

$$W(t) = \begin{pmatrix} y(t) & x(t) \\ y'(t) & x'(t) \end{pmatrix}, \tag{9}$$

does not vanish. Since $\det W(t)$ is independent of t (Abel), and since $y(t)$ can be multiplied by any non-vanishing constant, $y(t)$ can be chosen as follows: $\det W(t) \equiv 1$. Then the inverse of the matrix (9) is

$$W^{-1}(t) = \begin{pmatrix} x'(t) & -x(t) \\ -y'(t) & y(t) \end{pmatrix}. \tag{10}$$

Since $x(t)$ and $y(t)$ are solutions of the homogeneous equation (2), it is easily verified from (10) that, if $\phi(t)$ denotes the vector

$$\phi(t) = \int_0^t W^{-1}(u) \begin{pmatrix} 0 \\ g(u) \end{pmatrix} du, \tag{11}$$

then the first component of the vector $W(t)\phi(t)$ is a solution, $x = x_0(t)$, and the second component of (11) the derivative, $x_0'(t)$, of this solution, of the inhomogeneous equation, (3). (Needless to say, the differentiations occurring in this verification are equivalent to the proof of the rule for the variation of constants.) On the other hand, since either column of (9) is a vector the first component of which is a solution, and the second the derivative of this solution, of (2), the general solution of (2) results if an arbitrary constant vector is transformed by the matrix $W(t)$. Hence, if a and b denote the components of the arbitrary constant vector, then the first component of the vector

$$W(t)\phi(t) + W(t) \begin{pmatrix} a \\ b \end{pmatrix} \tag{12}$$

represents an arbitrary solution, and the second component the derivative of this solution, of (3).

By assumption, $f(t)$ is class (*). Since the definition of this class requires (6) not only for the given $x = x(t)$ but for $x = y(t)$ as well, it follows from (9) and (10) that all elements of both matrices $W(t)$, $W^{-1}(t)$ are $O(1)$ as $t \to \infty$. Consequently, both components of the vector (12) are of class (L^2) (for *fixed* a, b) if and only if both components of the vector

$$\phi(t) + \begin{pmatrix} a \\ b \end{pmatrix}$$

are of class (L^2) for the *same* values of the scalar constants a, b. For, on the one hand, these two vectors are transformed into one another by the matrices $W(t)$, $W^{-1}(t)$ and, on the other hand, the sum of two functions of class (L^2), as well as the product of two functions one of which is of class (L^2) and the other $O(1)$, is of class (L^2).

Accordingly, a solution of (3) is of class (L^2) *along with its derivative* if and only if the pair of integration constants a, b which belong to it by virtue of (12) is such as to render both components of the vector represented by the last formula line functions of class (L^2). However, since the assumptions made in (iii) with regard to $f(t)$ and $g(t)$ are assumed in (iv), the restriction just italicized, that concerning the derivative of a solution of (3), is made superfluous by (iii). On the other hand, (11) and (12) show that the vector in question is

$$\int_0^t \begin{pmatrix} x'(s) & -x(s) \\ -y'(s) & y(s) \end{pmatrix} \begin{pmatrix} 0 \\ g(s) \end{pmatrix} ds + \begin{pmatrix} a \\ b \end{pmatrix},$$

which means that the first component of the vector is

$$a - \int_0^t x(s) g(s) ds$$

and the second the negative of

$$-b - \int_0^t y(s) g(s) ds.$$

Since both of these expressions are of the form (8), the proof of (iv) is complete.

8. Under the first assumption of (i), the content of (i) can be amplified as follows:

(i*) *If $f(t)$ satisfies* (4) *and if a solution, $x(t)$, of* (2) *is of class* (L^2), *then both this solution and its derivative must tend to* 0, *whereas no solution linearly independent of this solution can remain bounded, as* $t \to \infty$.

The first of these three assertions, viz., (1), has been deduced above as a corollary of (i). In fact, only (5), rather than (4), was assumed in this deduction. The following proof of the second assertion, viz., of

$$(1^*) \qquad x'(t) \to 0 \text{ as } t \to \infty,$$

leaves it undecided whether or not (4) can again be relaxed to (5).

First, (i) implies that the product xx' is of class (L). It follows therefore from (4) that the same is true of f times xx'. In view of (2), this means that $x'x''$ is of class (L). Consequently, the indefinite integral of $x'x''$ must tend to a finite limit as $t \to \infty$. Since this indefinite integral is $\frac{1}{2}x'^2$, it follows that, if the assertion of the last formula line did not hold, $x'(t)$ could not be of class (L^2).

In order to prove the remaining assertion of the last italicized theorem, let $x = y(t)$ be a solution of (2) satisfying $y(t) = O(1)$. The assertion is that $y(x)$ and $x(t)$ must be linearly dependent.

What this assertion claims is the (identical) vanishing of the Wronskian, $xy' - yx'$, the latter being independent of t for any two solutions of (2). Hence, it is sufficient to show that $xy' - yx'$ tends to 0 as $t \to \infty$. But both x and x' tend to 0 as $t \to \infty$. Consequently, it is sufficient to ascertain that both y and y' are $O(1)$. But $y(t) = O(1)$ is assumed for the solution $x = y(t)$ of (2), and so (4) and (2) imply that $y''(t) = O(1)$. This completes the proof, since $y'(t) = O(1)$ is a necessity for every function $y(t)$ having a second derivative and satisfying both $y(t) = O(1)$ and $y''(t) = O(1)$, as $t \to \infty$ (Hadamard).

THE JOHNS HOPKINS UNIVERSITY.

IDENTITIES IN THE THEORY OF POWER SERIES.*

By Issai Schur.[1]

1. Introduction. Let

$$g(x) = 1 + a_1 x + a_2 x^2 + \cdots a_n x^n + \cdots \tag{1}$$

be a power series which converges in a certain neighborhood of $x = 0$. We write

$$[g(x)]^n = \sum_{\rho=0}^{\infty} a_{n\rho} x^\rho, \qquad (a_{n0} = 1); \qquad n = 0, \pm 1, \pm 2, \cdots, \tag{2}$$

and introduce the infinite set of series

$$\phi_n(x) = \sum_{\nu=\sigma}^{\infty} a_{n-\nu,\nu} x^\nu; \qquad n = 0, \pm 1, \pm 2, \cdots. \tag{3}$$

We wish to prove the following theorems:

I. The ratio

$$\frac{\phi_{n+1}(x)}{\phi_n(x)} = \psi(x) \tag{4}$$

is independent of n.

II. If we put $f(x) = xg(x)$, then

$$\psi[f(x)] = g(x). \tag{5}$$

III. For every $n = 0, \pm 1, \pm 2, \cdots$

$$f'(x)\phi_n[f(x)] = [g(x)]^{n+1}. \tag{6}$$

In order to establish the truth of these propositions, it suffices evidently to prove Theorem III. For, dividing two identities (6) for $n+1$ and n we get the identity (5) with $\psi(x) = \dfrac{\phi_{n+1}(x)}{\phi_n(x)}$; this proves further the inde-

* Received June 10, 1946.

[1] Died January 10th, 1941, Tel-Aviv, Palestine. This is the third posthumous paper of the deceased. The preceding two papers were: 1. "On Faber polynomials," *American Journal of Mathematics*, vol. 67 (1945), pp. 33-41. 2. "Ein Satz über quadratische Formen," *American Journal of Mathematics*, vol. 67 (1945), pp. 472-480. The present Note has been elaborated on the basis of a manuscript of Schur by M. Schiffer, Hebrew University, Jerusalem, in cooperation with Prof. M. Fekete.

pendence of $\psi(x)$ of the index n. Thus, Theorems I and II follow easily from Theorem III which will now be derived from Cauchy's Theorem.[2]

In fact, we have for small enough r

$$a_{n-\nu,\nu} = 1/2\pi i \int_{|\xi|=r} \frac{[g(\xi)]^{n-\nu}}{\xi^{\nu+1}}\, d\xi, \tag{7}$$

whence, in view of (3), for $|x|$ small enough

$$\phi_n(x) = 1/2\pi i \int_{|\xi|=r} \frac{[g(\xi)]^n}{\xi} \sum_{\nu=0}^{\infty} \frac{x^\nu}{\xi^\nu [g(\xi)]^\nu}\, d\xi \tag{8}$$

$$= 1/2\pi i \int_{|\xi|=r} \frac{g(\xi)^{n+1}}{\xi g(\xi) - x}\, d\xi.$$

Consider now the equation

$$x = yg(y) = f(y) \tag{9}$$

which possesses for small enough $|x|$ a unique solution y in the neighborhood of $y = 0$. Obviously, $\phi_n(x)$ is equal to the residue of the integrand of (8) at the point $\xi = y$. Hence, we get finally

$$\phi_n(x) = \phi_n[f(y)] = \frac{[g(y)]^{n+1}}{f'(y)}\,, \tag{10}$$

which proves Theorem III.

Let us write the identity (5) in the form

$$x\psi[f(x)] = f(x). \tag{5'}$$

This is equivalent to the theorem

IV. The function $y = f(x) = xg(x)$ has the inverse function

$$x = h(y) = y/\psi(y). \tag{11}$$

Another simple representation for the inverse function $h(y)$ is easily derived from (6). In fact, putting $n = -1$, we get

$$f'(x)[1 + a_{-2,1}f(x) + a_{-3,2}f(x)^2 + \cdots a_{-k-1,k}f(x)^k + \cdots] = 1. \tag{6'}$$

Integrating with respect to x between the limits 0 and x, we get identically in x

$$x = \sum_{\nu=1}^{\infty} (a_{-\nu,\nu-1}/\nu) f(x)^\nu. \tag{12}$$

V. The inverse function $x = h(y)$ of $y = f(x)$ has the development

---

[2] The following proof is due to Mr. George Schur.

$$x = h(y) = \sum_{\nu=1}^{\infty} (a_{-\nu,\nu-1}/\nu) y^{\nu}. \tag{12'}$$

The calculation of the coefficients $a_{-\nu,\mu}$ is very much facilitated by the following remark. The coefficients $a_{-1,\mu}$ in the development

$$[g(x)]^{-1} = \sum_{\mu=0}^{\infty} a_{-1,\mu} x^{\mu} \tag{13}$$

are given by the well known formula

$$a_{-1,\mu} = \sum_{\alpha_1+2\alpha_2+\ldots+\mu\alpha_\mu=\mu} (-1)^{\alpha_1+\alpha_2+\ldots+\alpha_\mu} \frac{(\alpha_1+\alpha_2+\cdots+\alpha_\mu)!}{\alpha_1!\alpha_2!\cdots\alpha_\mu!} a_1^{\alpha_1} a_2^{\alpha_2} \cdots a_\mu^{\alpha_\mu}. \tag{13'}$$

If, on the other hand, we differentiate identity (2) with respect to a_k and compare the coefficients of x^{μ} on both sides, we find

$$\partial a_{n\mu}/\partial a_1 = -\, n a_{n-1,\mu-k} \qquad (\mu = k, k+1, \cdots). \tag{14}$$

Thus, by means of (14), we may compute all coefficients $a_{-\nu,\mu}$ by differentiating the $a_{-1,\mu}$. In fact, we obtain

$$\partial a_{-n,\mu+1}/\partial a_1 = -\, n a_{-n-1,\mu} \tag{15}$$

whence, by iteration,

$$a_{-n,\mu} = \frac{(-1)^{n-1}}{(n-1)!} \frac{\partial^{n-1} a_{-1,\mu+n-1}}{\partial {a_1}^{n-1}} \tag{15'}$$

The coefficients in the development (12') have, therefore, the form

$$\frac{a_{-\nu,\nu-1}}{\nu} = \frac{(-1)^{\nu-1}}{\nu!} \frac{\partial^{\nu-1} a_{-1,2(\nu-1)}}{\partial {a_1}^{\nu-1}} \tag{15''}$$

and may be easily calculated from (13').

We further point out the following identity which results immediately from (11) and (12'):

$$\psi(x) \cdot \sum_{\nu=1}^{\infty} (a_{-\nu,\nu-1}/\nu) x^{\nu-1} \equiv 1. \tag{16}$$

The identity (12') is closely related to Lagrange's inversion formula. This formula permits to solve the equation

$$x = a + y g(x) \tag{17}$$

for any given function $g(x)$, analytic at $x = a$, for small enough y. If $\phi(x)$ is a function which is analytic at $x = a$, we obtain, by the formula mentioned above, the following development of $\phi(x)$ into a series in powers of y:

$$\phi(x) = \phi(a) + \sum_{\nu=1}^{\infty} y^\nu/\nu![(d^{\nu-1}/dx^{\nu-1})(\phi'(x)g(x)^\nu)]_{x=a} \tag{18}$$

which converges for small enough y.

Let us put in (18) $a = 0$, $\phi(x) = x$ and use our notations from (2). We obtain

$$x = \sum_{\nu=1}^{\infty} (a_{\nu,\nu-1}/\nu)y^\nu \tag{19}$$

as an inversion formula for

$$x = yg(x). \tag{19'}$$

If we replace $g(x)$ by $g(x) = g(x)^{-1}$ we have to put instead of $a_{\nu,\nu-1}$ the coefficient $a_{-\nu,\nu-1}$. Thus, the inversion of $y = xg(x)$ is given exactly by (12'). This identity is, therefore, a simple consequence of Lagrange's formula.

Next, we apply Lagrange's formula (18) with $a = 0$, $\phi(x) = x^k$ $(k = 1, 2, \cdots)$. We find

$$x^k/k = \sum_{\nu=\sigma}^{\infty} a_{k+\nu,\nu}(y^{k+\nu}/(k+\nu)). \tag{20}$$

Putting $x = yg(x)$, we immediately obtain

$$[g(x)]^k/k = \sum_{\nu=0}^{\infty} a_{-(k+\nu),\nu}(y^\nu/(k+\nu)). \tag{20'}$$

Let us replace here again $g(x)$ by $g(x)^{-1}$ and at the same time $a_{\nu,\mu}$ by $a_{-\nu,\mu}$. We obtain the identity

$$[g(x)]^{-k}/k = \sum_{\nu=0}^{\infty} a_{-(k+\nu),\nu}(y^\nu/(k+\nu)) \tag{21}$$

in x and y, if these variables are connected by the equation

$$y = xg(x) = f(x). \tag{21'}$$

Multiplying both sides of (21) by $y^k = x^k g(x)^k$, leads to

$$x^k/k = \sum_{\nu=0}^{\infty} a_{-(k+\nu),\nu}(y^{k+\nu}/(k+\nu)), \qquad (k = 1, 2, \cdots). \tag{21''}$$

Differentiating the last identity with respect to x, we obtain, in virtue of (3) and (21')

$$x^{k-1} = \sum_{\nu=0}^{\infty} a_{-(k+\nu),\nu} y^{k+\nu-1} \cdot f'(x) = y^{k-1} \cdot f'(x)\phi_{-k}[f(x)] \tag{22}$$

whence in view of (21')

$$[g(x)]^{-k+1} = f'(x)\phi_{-k}[f(x)]. \tag{22'}$$

Thus, we have derived identity (6) for n negative from Lagrange's formula.

2. Inverse Matrices. Lagrange's formula as well as our identity (6) may be proved by means of Cauchy's Theorem. The main task of this paper is to prove Theorem III in a purely arithmetic way and so open a new access to Lagrange's formula.

The decisive role in the establishment of the basic identities, used in this proof, is played by the following trivial

LEMMA. *Denote by* $[\psi(x)]_k$ *the coefficient of* x^k *in the development of the arbitrary analytic function* $\psi(x)$ *into powers of* x. *With this notation, the function* (1) *satisfies for* $k = 1, 2, \cdots$ *the equations*

$$[g(x)^k + xg(x)^{k-1}g'(x)]_k = 0, \tag{23}$$

$$[g(x)^{-k} + xg(x)^{-k-1}g'(x)]_k = 0. \tag{23'}$$

The truth of this lemma follows immediately from the fact that

$$g(x)^{k-1}g'(x) = (1/k)(g(x)^k)', \qquad g(x)^{-k-1}g'(x) = (-1/k)(g(x)^{-k})' \tag{24}$$

whence

$$[xg(x)^{k-1}g'(x)]_k = a_{k,k}, \qquad [xg(x)^{-k-1}g'(x)]_k = -a_{-k,k} \tag{24'}$$

which, in virtue of (2), proves (23) and (23').

By means of the preceding lemma, we prove now the following

THEOREM. *For every function* $g(x)$ *of the type* (6) *define the two matrices*

$$M = (a_{\mu,\mu-\nu}) = \begin{pmatrix} 1 & 0 & 0 \cdots \\ a_{21} & 1 & 0 \cdots \\ a_{32} & a_{31} & 1 \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \end{pmatrix}, \quad N = (a_{\nu,\mu-\nu}) = \begin{pmatrix} 1 & 0 & 0 \cdots \\ a_{11} & 1 & 0 \cdots \\ a_{12} & a_{21} & 1 \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \end{pmatrix} \tag{25}$$

with $a_{n\rho}$ vanishing for $\rho < 0$ and being defined by (2) for $\rho \geq 0$. Then, the inverse matrices are given by the formulae

$$M^{-1} = ((\mu/\nu)a_{-\nu,\mu-\nu}), \qquad N^{-1} = ((\nu/\mu)a_{-\mu,\mu-\nu}). \tag{26}$$

Proof. Obviously, all four matrices mentioned in the theorem are triangular matrices, the diagonals of which consist of units. To prove, therefore, that M and M^{-1} are in fact inverse matrices, it suffices to show that for $\mu > \nu \geq 1$ we have

$$J_{\mu\nu} = \sum_{\nu \leq \lambda \leq \mu} a_{\mu,\mu-\lambda} \cdot (\lambda/\nu) a_{-\nu,\lambda-\nu} = 0. \tag{27}$$

Now, it is easily verified that

$$(28) \qquad [g(x)^{\mu}\cdot(x^{\nu}g(x)^{-\nu})'/x^{\nu-1}]_{\mu-\nu}=\sum_{\nu\leq\lambda\leq\mu}a_{\mu,\mu-\lambda}\lambda a_{-\nu,\lambda-\nu}=\nu J_{\mu\nu}$$

and that, therefore,

$$(28') \quad \nu J_{\mu\nu}=[\nu g(x)^{\mu}(g(x)^{-\nu}-xg(x)^{-\nu-1}g'(x))]_{\mu-\nu}.$$
$$=\nu[g(x)^{\mu-\nu}-xg(x)^{\mu-\nu-1}g'(x)]_{\mu-\nu}=0$$

because of (23). Thus, we have proved (27).

In exactly the same way, we may prove for every $\mu>\nu\geq 1$ that

$$(29) \qquad J_{\mu\nu}=\sum_{\nu\leq\lambda\leq\mu}(\lambda/\mu)a_{-\mu,\mu-\lambda}\cdot a^{\nu}{}_{,\lambda-\nu}=0$$

which is equivalent to the matrix equation $N^{-1}\cdot N=1$. In fact, we have

$$(30) \qquad [g(x)^{-\mu}((x^{\nu}g(x)^{\nu})'/x^{\nu-1})]_{\mu-\nu}=\sum_{\nu\leq\lambda\leq\mu}\lambda a_{-\mu,\mu-\lambda}a_{\nu,\lambda-\nu}=\mu J_{\mu\nu}.$$

On the other hand, we may write the left hand expression in the form

$$(30') \quad [\nu g(x)^{-\mu}(g(x)^{\nu}+xg(x)^{\nu-1}g'(x))]_{\mu-\nu}.$$
$$=\nu[g(x)^{-(\mu-\nu)}+xg(x)^{-(\mu-\nu)-1}g'(x)]_{\mu-\nu}.$$

which vanishes in view of (23'). Hence, $y_{\mu\nu}=0$, which completes our proof.

We introduce a further matrix

$$(31) \qquad K=(\mu\delta_{\mu\nu})=\begin{pmatrix}1&0&0&0&\cdot\\0&2&0&0&\cdot\\0&0&3&0&\cdot\\\cdot&\cdot&\cdot&\cdot&\cdot\end{pmatrix}$$

and write $M=M(g)$, $N=N(g)$ in order to emphasize the dependence of these matrices on the function $g(x)$. Then, we obviously have in view of (25) and (26) the matrix equation

$$(32) \qquad M^{-1}(g)=KN(1/g)K^{-1},$$

since replacing g by $1/g$ means a change from $a_{\nu,\mu}$ to $a_{-\nu,\mu}$ in all matrices. Hence,

$$(33) \qquad K=M(g)\cdot K\cdot N(1/g)$$

or replacing g by $1/g$,

$$(33') \qquad K=M(1/g)\cdot K\cdot N(g)$$

i. e.,

$$(33'') \qquad N(g)^{-1}=K^{-1}M(1/g)K=((\nu/\mu)a_{-\mu,\mu-\nu}).$$

Thus, we see that the knowledge of M^{-1}, leading to the identity (32), provides a new way for calculating N^{-1}.

3. Lagrange's Inversion Formula. The matrix identities found in the last paragraph will be applied now to prove Lagrange's formula—and thus all its consequences—in a purely algebraic way.

First of all, we want to solve the equation

$$y = xg(x) \tag{21'}$$

by means of a series in powers of y

$$x = y + A_1 y^2 + A_2 y^3 + \cdots. \tag{34}$$

In order to determine all A_ν insert in (34) $y = xg(x)$ and compare on both sides the coefficient of x^m. Putting $\delta_{km} = \begin{cases} 1 & \text{for } m \begin{cases} = k \\ \neq k \end{cases} \\ 0 \end{cases}$ we get

$$\delta_{1m} = a_{1,m-1} + A_1 a_{2,m-2} + A_2 a_{3,m-3} + \cdots A_{m-1} a_{m,0}, \quad (m = 1, 2, \cdots). \tag{35}$$

$$\delta = \begin{pmatrix} \delta_{11} \\ \delta_{12} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \qquad A = \begin{pmatrix} 1 \\ A_1 \\ A_2 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}; \tag{35'}$$

in this notation (35) may be written as

$$\delta = N \cdot A \tag{35''}$$

which yields by inversion

$$A = N^{-1} \cdot \delta, \tag{36}$$

whence, in view of (26) and (35'),

$$A_\mu = (1/\mu + 1) a_{-(\mu+1),\mu}. \tag{37}$$

Thus, we have solved the equation (21') uniquely by the development

$$x = y + a_{-2,1}(y^2/2) + a_{-3,2}(y^3/3) + \cdots \tag{34'}$$

which is exactly the statement of Theorem V.

Next, let us deal with the more general question of determining the coefficients in the development

$$x^k = B_k y^k + B_{k+1} y^{k+1} + \cdots \tag{38}$$

where y and x are connected by the equation (21′). Comparing on both sides the coefficient of x^m, we obtain

$$\delta_{km} = B_k A_{k,m-k} + B_{k+1} A_{k+1,m-k-1} + \cdots + B_m a_{m,0}. \tag{39}$$

In order to express (39) in the form of a vector equation, we define the vectors

$$\delta_k = \begin{pmatrix} \delta_{k1} \\ \delta_{k2} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \qquad B = \begin{pmatrix} B_1 \\ B_2 \\ \cdot \\ \cdot \\ B_k \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \quad \text{with} \quad B_1 = \cdots B_{k-1} = 0,\ B_k = 1, \cdots. \tag{39′}$$

By means of these definitions we may write instead of (39)

$$\delta_k = N \cdot B \tag{39″}$$

whence, on multiplication by N^{-1},

$$B = N^{-1} \cdot \delta_k \qquad \text{i. e.,} \qquad B_\mu = (k/\mu) a_{-\mu,\mu-k}. \tag{40}$$

The development (38) has, therefore, the form

$$x^k = y^k + (k/(k+1)) a_{-(k+1),1} y^{k+1} + (k/(k+2)) a_{-(k+2),2}\, y^{k+2} + \cdots,\ k = 2, 3, \cdots. \tag{41}$$

The formula (41), valid in view of (34′) also for $k = 1$, is identical with (21″), i. e., with Lagrange's inversion formula for the equation $y = xg(x)$. We have derived here this important formula in a purely arithmetic way.

It is usual to give the inversion formula for the equation

$$x = yg(x). \tag{42}$$

We have already dealt in 1 with the connections of the corresponding results (compare formula (20)).

We have derived already in 1 the identity (6), for negative n, from equation (21) in a purely formal way. Hence, having proved (21) by arithmetic means only, we did the same for identity (6) in case of n negative. But this identity may be conceived as an identity between the coefficients of x^m on both sides, which are polynomials in n, a_1, a_2, $\cdots$. Therefore, these identities remain valid for n positive also.

We may even generalise this last result. We put

$$\gamma(x) = a_1 x + a_2 x^2 + \cdots, \qquad g(x) = 1 + \gamma(x) \tag{43}$$

and define for *every* real or complex κ

$$g(x)^\kappa = \sum_{\nu=0}^{\infty} \binom{\kappa}{\nu} \gamma(x)^\nu = \sum_{\nu=0}^{\infty} a_{\kappa,\nu} x^\nu, \quad \phi_\kappa(x) = \sum_{\nu=0}^{\infty} a_{\kappa-\nu,\nu} x^\nu. \tag{44}$$

Evidently, the $a_{\kappa-\nu,\nu}$ are still polynomials in $\kappa, a_1, a_2, \cdots$. Hence, the identity (6), being valid for $n = -1, -2, -3 \cdots$, remains valid for all complex values of n. Thus, we have proved the most general form of (6) by purely algebraic means.

Finally, we mention another group of identities. In analogy to $\phi_n(x)$ we define series

$$\psi_n(x) = \sum_{\nu=0}^{\infty} a_{n+\nu,\nu} x^\nu. \tag{45}$$

We emphasize the correspondence between ϕ_n, ψ_n and the initial function $g(x)$ by the notation $\phi_n(x; g)$ and $\psi_n(x; g)$. Then, we have obviously

$$\psi_n(x; g) = \phi_{-n}(x; (1/g)). \tag{46}$$

For every integer value of n we have, in view of (16),

$$g(x)^{n+1} = (xg(x))' \phi_n(xg(x); g). \tag{47}$$

Replacing $g(x)$ by $1/g(x)$ yields, in virtue of (46),

$$g(x)^{-(n+1)} = (x/g(x))' \phi_n(x/g(x); 1/g) = (x/g(x))' \psi_{-n}(x/g(x); g). \tag{47'}$$

Writing $-n$ instead of n yields clearly

$$g(x)^{n+1} = E(x) \psi_n(x/g(x); g), \qquad E(x) = g(x) - xg'(x). \tag{47''}$$

From (47'') it is obvious that

$$\psi(x) = \psi_{n+1}(x)/\psi_n(x) \tag{48}$$

is independent of n; but this result is also a consequence of definition (46) and Theorem I.

Further, we may write, in view of (47) and (47''),

$$g(x)^{n+1} = (g(x) + xg'(x)) \psi_n(xg(x) = (g(x) - xg'(x)) \psi_n(x/g(x)), \tag{49}$$

whence the result that

$$\psi_n(x/g(x))/\phi^n(xg(x)) = g(x) + xg'(x)/g(x) - xg'(x) \tag{49'}$$

is independent of n.

4. Another Inversion Formula. In this paragraph we shall consider series of the type

$$(50)\quad w = z + a_1 + a_2/z + a_3/z^2 + \cdots = zg(1/z),$$
$$g(x) = 1 + a_1 x + a_2 x^2 + \cdots.$$

We wish to invert (50) into the form

$$(51)\qquad z = w + A_1 + A_2/w + \cdots.$$

In order to determine the coefficients A_ν, put $x = 1/z$ and consider the identity in z:

$$(51')\qquad z = zg(x) + A_1 + A_2 x/g(x) + A_3 x^2/g(x)^2 + \cdots.$$

Comparing the coefficients of x^ν on both sides of this identity yields

$$(52)\qquad A_1 = -a_1, \qquad a_\nu + \sum_{\rho=1}^{\nu-1} A_{\rho+1} a_{-\rho,\nu-\rho-1} = 0 \qquad (\nu = 2, 3, \cdots).$$

We introduce the vectors

$$(52')\qquad C = \begin{bmatrix} -a_2 \\ -a_3 \\ -a_4 \\ . \\ . \\ . \\ . \end{bmatrix}, \qquad D = \begin{bmatrix} A_2 \\ A_3 \\ . \\ . \\ . \\ . \\ . \end{bmatrix};$$

then, the system of equations of (52) may be written, in view of definition (25), in a single vector formula

$$(52'')\qquad C = N(1/g) \cdot D,$$

whence, in virtue of (26),

$$(53)\qquad D = N(1/g)^{-1} \cdot C, \quad \text{i. e.,} \quad \mu A_{\mu+1} = -\sum_{\lambda=1}^{\mu} \lambda a_{\mu,\mu-\lambda} \cdot a_{\lambda+1}.$$

We introduce the series

$$(54)\qquad \phi(x) = a_2 x^2 + 2a_3 x^3 + \cdots + n a_{n+1} x^{n+1} + \cdots.$$

We have by (53) evidently

$$(55)\qquad [g(x)^n \cdot \phi(x)]_{n+1.} = \sum_{\lambda=1}^{n} \lambda a_{n,n-\lambda} \cdot a_{\lambda+1} = -nA_{n+1}.$$

On the other hand,

(56) $$g(x)+\phi(x)-1=a_1x+2a_2x^2+3a_3x^3+\cdots=xg'(x);$$

hence, in virtue of (55) and (56),

(57) $$-nA_{n+1}=[xg(x)^n\cdot g'(x)-g(x)^{n+1}+g(x)^n]_{n+1}.$$

Now, we apply the lemma of **2**; in virtue of (23) with $k=n+1$ we have

(57′) $$-nA_{n+1}=[g(x)^n]_{n+1.}=a_{n,n+1}.$$

We find, therefore, for (51) the simple expression

(58) $$z=w-a_1-a_{1,2}/1\cdot(1/w)-a_{2,3}/2\cdot(1/w^2)-\cdots-a_{n,n+1}/n\cdot(1/w^n)-$$

If we write $x=1/z$, $y=1/w$, we may put (50) into the form

(42′) $$x=yg(x).$$

This allows the following interpretation of (58):

(59) $$-1/x=-1/y+a_1+(a_{1,2}/1)y+(a_{2,3}/2)y^2 +\cdots+(a_{n,n+1}/n)y^n+\cdots,$$

i. e., this series is a generalization of Lagrange's inversion formula (20) for $k=-1$. One might be tempted to apply directly (20) with $k=-1$; but in this formula appears the term $a_{k+1,1}/(k+1)$ which is undefined for $k=-1$. If, however, we replace this meaningless term by a_1 the formula (20) with $k=-1$ is the same, term by term, as (59).

Let us write (20) with $k=1$; we obtain

(60) $$x=y+(a_{2,1}/2)y^2+(a_{3,2}/3)y^3+\cdots+(a_{n,n-1}/n)y^n+\cdots.$$

Multiplying (59) by (60) we obtain the identity

(61) $$(1-a_1y-(a_{1,2}/1)y^2-(a_{2,3}/2)y^3-\cdots)\\ (1+(a_{2,1}/2)y+(a_{3,2}/3)y^2+\cdots)=1.$$

There arises now the question of generalizing Lagrange's formula (20) for general k. We proceed in an analogous way as in **3**. We solve equation (42) by the development (60):

(62) $$x=yh(y)=y(1+(a_{2,1}/2)y+(a_{3,2}/3)y^2+\cdots).$$

We introduce this development into the formula (20), valid for positive integer values of k. After dividing by y^k we obtain

(63) $$h(y)^k/k=\sum_{\nu=0}^{\infty}(a_{k+\nu,\nu}/k+\nu)y^\nu.$$

If we write

$$h(y)^n = \sum_{\nu=0}^{\infty} b_{n,\nu} y^\nu \qquad (n = 0, \pm 1, \pm 2, \cdots) \tag{64}$$

we get, by comparing the coefficients of like powers of y^m on both sides of (63),

$$b_{k,m}/k = a_{k+m,m}/k + m \qquad (k = 1, 2, 3 \cdots ; m = 0, 1, 2, \cdots). \tag{65}$$

By definitions (2), (62) and (64), the $a_{k+m,m}$ and $b_{k,m}$ are, for m fixed, polynomial functions of $k, a_1, a_2, \cdots$. Therefore, on each side of (65) stands a rational function of k and both functions coincide for positive integer values of k. Hence, these functions are identical and the equality (65) holds for *every* (complex) value of k. Formula (20) is, therefore, also true for every k, except for negative integers for which one of its terms becomes meaningless.

In order to generalize (20) for the case of negative integers too, we have only to replace the undefined quotient $a_{k+n,n}/(k+n)$ in the case $k = -n$ by the limit

$$\lim_{k \to -n} a_{k+n,n}/(k+n).$$

We have obviously

$$\lim_{l=0} \frac{g(x)^l - 1}{l} = \log\, g(x), \tag{66}$$

whence, by comparing the coefficients of x^m on both sides,

$$\lim_{l=0} a_{l,m}/l = [\log\, g(x)]_m. \qquad (m = 1, 2, \cdots). \tag{67}$$

Hence, in the case $k = -n$, we have to replace in (20) the term $a_{k+n,n}/(k+n)$ by $[\log g(n)]_n$.

If we put

$$\log\, g(x) = \sum_{\nu=1}^{\infty} \gamma_\nu x^\nu \tag{68}$$

we may write Lagrange's inversion formula (20) for $k = -n$ in the form

$$\begin{aligned} \frac{x^{-n}}{-n} = \frac{y^{-n}}{-n} &+ (a_{-(n-1),1}/-(n-1))y^{-(n-1)} \\ &+ \cdots (a_{-1,n-1}/-1)y^{-1} + \gamma_n + (a_{1,n+1}/1)y + \cdots. \end{aligned} \tag{69}$$

In a similar way, we obtain a new identity by adding to both sides of (63) the term $-1/k$ and then passing to the limit $k = 0$. We obtain

$$\log\, h(y) = \sum_{\nu=1}^{\infty} (a_{\nu,\nu}/\nu) y^\nu. \tag{70}$$

On the other hand, we have in view of $x = yg(x)$ and $x = yh(y)$ obviously $h(y) = g(x)$ and thus

$$\log g(x) = \sum_{\nu=1}^{\infty} (a_{\nu,\nu}/\nu) y^{\nu}. \tag{70'}$$

Finally, we apply the formula (69) in the case of series (50). We put, as before, $x = 1/z$, $y = 1/w$ which transforms equation (50) into (42) and vice versa. Instead of (69), we obtain, for positive integers n,

$$z^n/n = w^n/n + (a_{-(n-1),1}/(n-1))w^{(n-1)} + \cdots + (a_{-1,n-1}/1)w - \gamma_n - a_{1,n+1}/1 \cdot (1/w) - a_{2,n+2}/2 \cdot (1/w^2) - \cdots. \tag{71}$$

Consider now the polynomial

$$P_n(w) = n[w^n/n + (a_{-(n-1),1}/(n-1))w^{n-1} + \cdots + (a_{-1,n-1}/1)w - \gamma_n]; \tag{72}$$

it has the remarkable property that by introducing $w = zg(1/z) = f(z)$ we get a development

$$P_n(f(z)) = z^n + \alpha_1/z + \alpha_2/z^2 + \cdots, \tag{72}$$

with z^n as the only non-negative power of z. By this characteristic property the polynomial $P_n(w)$ is associated uniquely with the series $w = f(z)$ and is called the n-th Faber polynomial of $f(z)$. It plays an important role in the theory of conformal representation.[3]

On comparing the developments (68) and (70') we obtain, in view of (20), the equality

$$\sum_{k=1}^{\infty} \gamma_k x^k = \sum_{k=1}^{\infty} k\gamma_k \sum_{\nu=0}^{\infty} (a_{k+\nu,\nu}) y^{k+\nu}/k + \nu = \sum_{\nu=1}^{\infty} (a_{\nu,\nu}/\nu) y^{\nu}. \tag{73}$$

Comparing the coefficients of equal powers of y on both sides yields

$$\sum_{\nu=0}^{m-1} (m-\nu)\gamma_{m-\nu} a_{m,\nu} = a_{m,m} \qquad (m = 1, 2, \cdots), \tag{74}$$

an interesting relation between the $a_{\mu\nu}$.

---

[3] Compare: Schur, I., "On Faber Polynomials," *American Journal of Mathematics*, vol. 67 (1945), pp. 33-41.

A NOTE ON DIVISION RINGS.*

By N. JACOBSON.

In this note we consider division subrings Φ of a division ring P such that the left dimensionality $(P : \Phi)_l$ is finite. If $\mathfrak{L}$ denotes the ring of linear transformations in P regarded as a vector space over the set Φ_l of left multiplications by elements of Φ then clearly $\mathfrak{L}$ contains the subring P_r of right multiplications by elements of P. The relation $(\mathfrak{L} : P_r)_r = (P : \Phi)_l$ can be established (Theorem 1). Thus $\mathfrak{L}$ has finite right dimensionality over P_r. The principal result of this paper is that the converse of these results holds, namely, if $\mathfrak{A}$ is any ring of endomorphisms of the additive group of P such that 1) $\mathfrak{A}$ contains P_r and 2) $(\mathfrak{A} : P_r)_r$ is finite, then $\mathfrak{A}$ coincides with a ring $\mathfrak{L}$ of linear transformations of P over Φ_l where Φ is a division subring of P such that P over Φ_l is finite. A complete reciprocity is established between the rings $\mathfrak{A}$ (or $\mathfrak{L}$) and the subrings Φ of P.

There are a number of applications of our results. One of these, the theory of automorphisms in a division ring, is developed here. In this connection we generalize the concept of a closed group of automorphisms introduced by Emmy Noether [1] for groups of inner automorphisms and we establish a Galois correspondence between closed groups that satisfy a certain finiteness condition and their division subrings of invariants. The theorem of Noether [2]-Artin and Whaples [3] on commutators of division subalgebras containing the center Γ and our own results on finite groups of outer automorphisms are contained as special cases of the theorems derived here.

1. Let P be a division ring and Φ a division subring of P of *finite left index m*. By this we mean that the left dimensionality $(P : \Phi)_l = m < \infty$. Here we are regarding P as a left vector space (or module) over Φ or, what amounts to the same thing, as a group relative to the set Φ_l of left multiplications $\xi \to \xi\alpha_l \equiv \alpha\xi$, ξ in P, α in Φ. Let $\mathfrak{L}$ denote the ring of linear transformations (l. t.) of P over Φ_l. Thus $A \in \mathfrak{L}$ if and only if A is an endomorphism of the additive group of P and

* Received June 28, 1946.

[1] [5], p. 529.

[2] [5], p. 528.

[3] [1], p. 104.

$$(\alpha\xi)A = \alpha(\xi A)$$

for all $\alpha \in \Phi$ and all $\xi \in \mathbf{P}$. Evidently $\mathfrak{L}$ contains the set $\mathbf{P}_r$ of right multiplications $\xi \to \xi\rho_r \equiv \xi\rho$ in $\mathbf{P}$. Since $1 \in \mathbf{P}_r$, $\mathfrak{L}$ may be regarded as a (right) vector space over $\mathbf{P}_r$. When this is done we obtain the following theorem:

THEOREM 1. *The dimensionalities* $(\mathfrak{L} : \mathbf{P}_r)_r$ *and* $(\mathbf{P} : \Phi)_l$ *are equal.*[4]

Let $\xi_1, \xi_2, \cdots, \xi_m$ be a left basis of $\mathbf{P}$ over Φ. Then any l. t. in $\mathbf{P}$ over Φ_l is completely determined by its effect on the ξ_i. Now define C_j, $j = 1, 2, \cdots, m$, to be the l. t. such that

$$\xi_j C_j = 1 \quad \text{and} \quad \xi_i C_j = 0 \text{ if } i \neq j.$$

Then $\Sigma C_j \eta_{jr}$ maps ξ_i into η_i. Hence $\Sigma C_j \eta_{jr} = 0$ implies that all the $\eta_j = 0$. Thus the C's are right independent relative to $\mathbf{P}_r$. On the other hand if A is any l. t. in $\mathbf{P}$ over Φ_l, let $\xi_i A = \eta_i$. Then A coincides with $\Sigma C_j \eta_{jr}$. Hence the C's form a right basis for $\mathfrak{L}$ over $\mathbf{P}_r$.

2. We assume now that $\mathfrak{A}$ is any ring of endomorphisms acting in the additive group of $\mathbf{P}$ that 1) contains $\mathbf{P}_r$ and 2) has finite dimensionality $(\mathfrak{A} : \mathbf{P}_r)_r$. Let C be any endomorphism in $\mathbf{P}$ that commutes with every $A \in \mathfrak{A}$. Then since C commutes with every right multiplication, $C = \gamma_l$ a left multiplication.[5] The totality of γ obtained in this way is a division subring Φ of $\mathbf{P}$ that we shall call the division ring of $\mathfrak{A}$-*invariants.*

THEOREM 2. *Let* $\mathfrak{A}$ *be a ring of endomorphisms of the additive group of* $\mathbf{P}$ *such that* 1) $\mathfrak{A} \supseteq \mathbf{P}_r$ *and* 2) $(\mathfrak{A} : \mathbf{P}_r)_r = m < \infty$ *and let* Φ *be the division subring of* $\mathfrak{A}$-*invariants. Then* $(\mathbf{P} : \Phi)_l = m$ *and* $\mathfrak{A} = \mathfrak{L}$ *the complete ring of* l. t. *of* $\mathbf{P}$ *over* Φ_l.

Since $\mathfrak{A} \supseteq \mathbf{P}_r$, $\mathfrak{A}$ is an irreducible ring of endomorphisms in $\mathbf{P}$. Since $(\mathfrak{A} : \mathbf{P}_r)_r < \infty$, $\mathfrak{A}$ satisfies the descending chain condition for right ideals. Now Φ_l is the division ring of endomorphisms commutative with the elements of the irreducible ring $\mathfrak{A}$. It follows from known results on irreducible rings of endomorphisms that $(\mathbf{P} : \Phi)_l < \infty$ and that $\mathfrak{A} = \mathfrak{L}$ the complete ring of l. t. of $\mathbf{P}$ over Φ_l.[6] By Theorem 1 $(\mathbf{P} : \Phi)_l = (\mathfrak{A} : \mathbf{P}_r)_r$.

To complete the reciprocity between Φ and $\mathfrak{L}$ we require the following theorem:

---

[4] The present proof of this theorem, which is simpler than my original one, was communicated to me by R. Brauer.

[5] [2], p. 54.

[6] [4], in particular Theorems 6 and 3.

THEOREM 3. *Let Φ be a division subring of finite left index* P *and let* $\mathfrak{L}$ *be the ring of* l. t. *in* P *over* Φ_l. *Then reciprocally* Φ_l *is the complete set of endomorphisms in* P *that commute with all the endomorphisms belonging to* $\mathfrak{L}$.

This result is well known for arbitrary vector spaces.[7] It can also be obtained as a consequence of Theorems 1 and 2. For let Ψ be the division ring of $\mathfrak{L}$-invariants. Then by Theorem 2, $(\mathrm{P} : \Psi)_l = (\mathfrak{L} : \mathrm{P}_r)_r$. On the other hand $\Psi \supseteq \Phi$ and $(\mathrm{P} : \Phi)_l = (\mathfrak{L} : \mathrm{P}_r)_r$ by Theorem 1. Hence $\Psi = \Phi$.

If $(\mathrm{P} : \Phi)_l < \infty$ and $\mathfrak{L}$ is the ring of l. t. of P over Φ_l our results establish a $(1-1)$ correspondence between the division rings Ψ between P and Φ and the rings $\mathfrak{B}$ between P_r and $\mathfrak{L}$. If Ψ is given then $\mathfrak{B}$ is the ring of l. t. of P over Ψ_l. Clearly $\mathrm{P}_r \subseteq \mathfrak{B} \subseteq \mathfrak{L}$. On the other hand if $\mathfrak{B}$ is a ring between P_r and $\mathfrak{L}$ then the set Ψ of $\mathfrak{B}$-invariants is a division subring containing Φ. These two correspondences are inverses of each other.

All of the above results can, of course, be duplicated for division subrings of finite right index. In the remainder of this note we consider some applications of these theorems to the theory of automorphisms in a division ring.

3. Let A be an automorphism in P. Then A is an endomorphism of the additive group of P and $(\xi\eta)A = (\xi A)(\eta A)$. Thus we have the following fundamental commutation rules governing automorphisms and multiplications:

$$\eta_r A = A(\eta A)_r \qquad \xi_l A = A(\xi A)_l. \tag{2}$$

We consider now the possible forms of linear relations with coefficients that are multiplications that can hold for automorphisms in P. First let $I_1, I_2, \cdots I_h$ be inner automorphisms, say, $\xi I_j = \tau_j \xi \tau_j^{-1}$. Then $I_j = \tau_{jl}\tau_{jr}^{-1}$ if τ_{jl} is the left multiplication determined by τ_j and τ_{jr} is the right multiplication determined by τ_j. Suppose that the τ's are linearly dependent over Γ where Γ is the center of P. Then there exist γ_i not all 0 such that $\Sigma\tau_i\gamma_i = 0$. Hence $\Sigma I_i\tau_{ir}\gamma_{ir} = \Sigma\tau_{il}\gamma_{ir} = \Sigma\tau_{il}\gamma_{il} = 0$. Since not every $\tau_{jr}\gamma_{jr}$ equals 0 this shows that the I_j are linearly dependent over P_r.

Conversely suppose that we have a relation $\Sigma I_i\eta_{ir} = 0$ where not every $\eta_{jr} = 0$. Then $\Sigma\tau_{il}\mu_{ir} = 0$ where $\mu_j = \tau_j^{-1}\eta_j$ so that not every $\mu_j = 0$. We assert that this implies that the τ_j are dependent over Γ. This is the following well-known lemma:

LEMMA 1. *If $\tau_1, \tau_2, \cdots \tau_h$ are elements of* P *that are linearly independent over* Γ *then the left multiplications $\tau_{1l}, \tau_{2l}, \cdots, \tau_{hl}$ are linearly independent over* P_r.

---

[7] [2], p. 22.

For the sake of completeness we give the proof. If the τ_l's are dependent over P_r there exists a shortest relation which we can suppose to be of the form

$$\tau_{1l}\mu_{1r} + \tau_{2l}\mu_{2r} + \cdots + \tau_{sl}\mu_{sr} = 0 \tag{3}$$

where all the μ_i are different from 0. We may assume also that $\mu_{1r} = 1_r$. Since the τ_j are independent over Γ, one of the μ's, say μ_2, is not in Γ. Hence there exists a ν such that $\mu_2\nu - \nu\mu_2 \neq 0$. Multiplication of (3) on the left and on the right by ν_r followed by subtraction of the resulting relations gives a new relation shorter than (3). This contradiction proves the result.

We suppose next that $A_1, A_2, \cdots, A_u$ are automorphisms that are in distinct cosets of the subgroup of inner automorphisms. Then no $A_iA_j^{-1}$, $i \neq j$, is inner. Let $\mathfrak{U}$ denote the ring of endomorphisms generated by the left and the right multiplications. The elements of $\mathfrak{U}$ have the form $U = \tau_{1l}\mu_{1r} + \tau_{2l}\mu_{2r} + \cdots$. We wish to prove the following lemma.

LEMMA 2. *If $A_1, A_2, \cdots, A_u$ are automorphisms such that no $A_iA_j^{-1}$, $i \neq j$, is inner, then a relation $\Sigma A_iU_i = 0$ where the $U_j \in \mathfrak{U}$ can hold only if every $U_j = 0$.*

We may choose elements $\tau_1, \tau_2, \cdots, \tau_h$ such that the left multiplications τ_{il} are independent over P_r and such that every $U_i = \Sigma\tau_{jl}\mu_{jir}$. Then our relation reads $\Sigma A_i\tau_{jl}\mu_{jir} = 0$. We shall show that every $\mu_{ji} = 0$. For otherwise we may suppose that $\Sigma A_i\tau_{jl}\mu_{jir} = 0$ is a shortest linear relation connecting the hu elements $A_i\tau_{jl}$ in which the coefficients $\mu_{ji_r} \neq 0$. We may suppose also that $\mu_{11} = 1$. We shall show that no A other than A_1 occurs in this relation. For suppose that

$$A_1\tau_{1l}\mu_{11r} + A_2\tau_{kl}\mu_{k2r} + \cdots = 0. \tag{4}$$

Multiply (4) on the left by ξ_r and on the right by $(\xi A_1)_r$ and subtract the resulting relations. This yields

$$A_2\tau_{kl}[(\xi A_2)_r\mu_{k2r} - \mu_{k2r}(\xi A_1)_r] + \cdots = 0. \tag{5}$$

If the bracket is 0 for all ξ

$$\xi A_2 = \mu_{k2}(\xi A_1)\mu_{k2}^{-1}$$

contrary to the assumption that A_1 and A_2 do not differ by an inner automorphism. Hence there is a ξ such that the bracket is not 0 and this gives a shorter relation than (4). Thus (4) must have the form

$$A_1(\tau_{1l}\mu_{11r} + \cdots) = 0.$$

Since A_1 is an automorphism this implies that $\tau_{1l}\mu_{11r} + \cdots = 0$ contrary to the assumption that the τ_{jl} are linearly independent over $\mathbf{P}_r$.

Suppose finally that we have an arbitrary linear relation with coefficients that are in $\mathbf{P}_r$ connecting a finite set of automorphisms. We may write this in the form $\Sigma A_i I_j \mu_{ji r} = 0$ where the I_j are inner and the A_i are in different cosets relative to the subgroup of inner automorphisms. If $I_j = \tau_{jl}\tau_{jr}^{-1}$, $\Sigma A_i \tau_{jl}\tau_{jr}^{-1}\mu_{ji r} = 0$. Hence by Lemma 2, $\Sigma I_j \mu_{ji r} = \Sigma \tau_{jl}\tau_{jr}^{-1}\mu_{ji r} = 0$. By Lemma 1, either all the μ's are 0 or the τ_j are linearly dependent over Γ.

4. An element ϕ of $\mathbf{P}$ is *invariant* under an automorphism A if $\phi A = \phi$. By (2) ϕ has this property if and only if the left (right) multiplication $\phi_l(\phi_r)$ commutes with A. If G is a group of automorphisms we shall call ϕ a *G-invariant* if ϕ is invariant under every $A \in G$. The totality of these elements is a division subring Φ of $\mathbf{P}$. Let $\mathfrak{A} = G\mathbf{P}_r$ denote the totality of endomorphisms in the additive group of $\mathbf{P}$ that have the form $\Sigma A_i \eta_{ir}$ where the $A_i \in G$. Clearly $\mathfrak{A}$ is closed under addition and subtraction and by the first equation of (2) $\mathfrak{A}$ is closed under multiplication. Hence $\mathfrak{A}$ is a subring of the ring of endomorphisms of $\mathbf{P}$. Since G contains the identity, $\mathfrak{A}$ contains $\mathbf{P}_r$. Thus an endomorphism commutes with every element of $\mathfrak{A}$ if and only if it is a left multiplication that commutes with every $A \in G$. It follows that the division ring of G-invariants coincides with the set of $\mathfrak{A}$-invariants in the sense defined above.

Let H be the invariant subgroup of inner automorphisms contained in G and let $I_1 = \tau_{1l}\tau_{1r}^{-1}$ and $I_2 = \tau_{2l}\tau_{2r}^{-1}$ belong to H. Then clearly $I_3 = I_1 I_2 = \tau_{3l}\tau_{3r}^{-1}$ where $\tau_3 = \tau_1\tau_2$. An element ϕ is an I_1 and an I_2-invariant if and only if it commutes with τ_1 and with τ_2. If this is so, ϕ also commutes with every $\tau = \tau_1\gamma_1 + \tau_2\gamma_2$, γ_i in Γ. Hence we do not change the set of G-invariants if we adjoin to G the inner automorphisms by the element τ obtained in this way. This remark leads us to restrict our attention to the study of groups G that are *closed* in the sense that if $I_1 = \tau_{1l}\tau_{1r}^{-1}$ and $I_2 = \tau_{2l}\tau_{2r}^{-1}$ are in G then so is $I = \tau_l\tau_r^{-1}$, where $\tau = \tau_1\gamma_1 + \tau_2\gamma_2$ and the γ_i are arbitrary in Γ. It is readily seen that G is closed if and only if the subgroup H is precisely the set of inner automorphisms that are defined by the non-zero elements of the division subalgebra $\mathbf{A}$ of $\mathbf{P}$ that contains the center Γ. The algebra $\mathbf{A}$, which is uniquely determined, will be referred to as the division subalgebra *associated* with the closed group G (or H).

We shall say that a closed group G has finite reduced order if 1) $(\mathbf{A} : \Gamma)$ is finite and 2) H has finite index in G. The product of these two numbers will be called *the reduced order* of G. Our aim in the remainder

of this note is the establishment of a Galois correspondence between closed groups of automorphisms of finite reduced order and their division rings of invariants. We prove first the following theorem:

THEOREM 4. *Let G be a closed group of automorphisms of finite reduced order in the division ring* P *and let $\mathfrak{A} = G\mathrm{P}_r$ be the ring of endomorphisms of the form $\Sigma A_i \eta_i r$, A_i in G. Then $(\mathfrak{A} : \mathrm{P}_r)_r =$ reduced order of G.*

Let $A_1, \cdots, A_u$ be representatives of the cosets of H in G and let $I_1, \cdots, I_h$ be inner automorphisms such that if $I_j = \tau_{j1}\tau_{jr}^{-1}$ then the elements τ_j form a basis for the subalgebra A associated with G. Then any automorphism in G has the form $A_k I$ where I is in H. If $I = \tau_1\tau_r^{-1}$, $\tau = \Sigma\tau_i\gamma_i$. As we have seen this implies that $I = \Sigma I_j \eta_{jr}$ where $\eta_j = \tau_j\gamma_j\tau^{-1}$. Hence $A_k I$ is a linear combination of the automorphisms $A_k I_1, A_k I_2, \cdots, A_k I_h$. Consequently the hu automorphisms $A_i I_j$ are generators over P_r of $\mathfrak{A}$. Now if $\Sigma A_i I_j \mu_{ji} r = 0$ then $\Sigma I_j \mu_{ji} r = 0$ by Lemma 2. Moreover $\Sigma I_j \mu_{ji} r = 0$ and the independence of the τ_j over Γ implies that all the $\mu_{ji} = 0$. Thus the $A_i I_j$ are independent over P_r. Hence $(\mathfrak{A} : \mathrm{P}_r)_r = hu$.

5. We may now establish the Galois correspondence between groups of automorphisms and their division rings of invariants. The basic results are contained in Theorems 5 and 8.

THEOREM 5. *Let G be a closed group of automorphisms of finite reduced order in the division ring* P *and let Φ be the division subring of G-invariants. Then $(\mathrm{P} : \Phi)_l =$ reduced order of $G = (\mathrm{P} : \Phi)_r$ and any automorphism of* P *that leaves the elements of Φ invariant belongs to G.*

Let $\mathfrak{A} = G\mathrm{P}_r$. Then by Theorem 4, $(\mathfrak{A} : \mathrm{P}_r)_r =$ reduced order of G. On the other hand we know that Φ is the division ring of $\mathfrak{A}$-invariants. Hence by Theorem 2, $(\mathrm{P} : \Phi)_l = (\mathfrak{A} : \mathrm{P}_r)_r$. Hence $(\mathrm{P} : \Phi)_l =$ reduced order of G. The equality $(\mathrm{P} : \Phi)_r =$ reduced order of G follows by symmetry since the concept of reduced order is a "two-sided" one. Suppose now that A is any automorphism in P with the property that $\phi A = \phi$ for all $\phi \in \Phi$. Then $A\phi_l = \phi_l A$ and A is a linear transformation in P over Φ_l. Hence by Theorem 2, $A \in \mathfrak{A}$. Thus $A = \Sigma A_i I_j \mu_{ji} r$ where the A_i and I_j are determined as in the proof of Theorem 4. By Lemma 2 this relation must have the form $A \equiv A_1 I = A_1 \Sigma I_j \mu_{jr}$. If $I = \tau_1\tau_r^{-1}$, Lemma 1 implies that τ is Γ-dependent on the τ_j, $j = 1, \cdots, h$. Hence $I \in H$ and $A = A_1 I \in G$.

The special case of this theorem in which $G = H$ consists of inner automorphisms is due to Artin and Whaples.[8] We note also that if G is a finite

[8] [1], p. 104.

group of automorphisms A_i such that each $A_i \neq 1$ is outer, then G is closed and of finite order. Hence Theorem 5 applies in this case too. The result thus obtained has been given in a previous paper of the author's.[9] A more general result can be obtained by combining Theorem 5 with the following theorem:

THEOREM 6. *Let F be any finite group of automorphisms and let J be its subgroup of inner automorphisms. Then the closed group G generated by F has finite reduced order. If $J = \{\tau_{1l}\tau_{1r}^{-1}, \cdots, \tau_{ml}\tau_{mr}^{-1}\}$ then the reduced order of G is the product of the order of the set $(\tau_1, \cdots, \tau_m)$ over Γ by the index of J in F.*

Since J is closed under multiplication $\tau_i\tau_j = \tau_k\gamma_{ij}$, γ_{ij} in Γ. If A is any automorphism and I is the inner automorphism $\tau_l\tau_r^{-1}$ then $A^{-1}IA$ is the inner automorphism $(\tau A)_l(\tau A)_r^{-1}$. Hence if $A \in F$, $(\tau_i A)_l(\tau_i A)_r^{-1} \in J$. It follows that $\tau_i A = \tau_h\delta_h$, δ_h in Γ. Now let $\mathbf{A}$ denote the totality of linear combinations $\Sigma\tau_i\gamma_i$, γ_i in Γ. Since the product of any two τ's is a multiple of a τ, $\mathbf{A}$ is a subalgebra. Clearly $\mathbf{A}$ has a finite basis and the dimensionality $(\mathbf{A} : \Gamma)$ is the order of the set $(\tau_1, \cdots, \tau_m)$ over Γ. Since $\mathbf{A}$ has a finite basis, $\mathbf{A}$ is a division algebra. Let H denote the group of inner automorphisms determined by the elements of $\mathbf{A}$. Evidently H is contained in the closed group G generated by F. If $A \in F$ and $I = \tau_l\tau_r^{-1} \in H$ where $\tau = \Sigma\tau_i\gamma_i$ then $A^{-1}IA$ is the inner automorphism determined by the element $\Sigma(\tau_i A)(\gamma_i A)$. Since $\tau_i A = \tau_h\delta_h$, δ_h in Γ and $\gamma_i A \in \Gamma$, $A^{-1}IA = I'$ is in H. It follows that the totality G' of automorphisms of the form AI, A in F and I in H is a group. For the product $(A_1I_1)(A_2I_2) = A_1A_2(A_2^{-1}I_1A_2)I_2$ is in this set and $(AI)^{-1} = I^{-1}A^{-1} = A^{-1}(AI^{-1}A^{-1})$ is in the set. If AI is inner, so is A. Hence A is in J. Thus $AI \in H$. This shows that H is the group of inner automorphisms contained in G'. Since H is a group determined by the elements of an algebra, G' is closed. Hence $G' = G$ the smallest closed group containing F. Since the elements of G have the form AI, A in F, it is easy to see that the index of H in G is the same as that of J in F. It follows that G is of finite reduced order and that its reduced order is the product of $(\mathbf{A} : \Gamma)$ by the index of J in F.

6. Let G be an arbitrary closed group of finite reduced order and let $\mathfrak{A} = G\mathbf{P}_r$. We wish to establish a $(1 - 1)$ correspondence between the closed subgroups K of G and the subrings $\mathfrak{L}$ of $\mathfrak{A}$ that contain $\mathbf{P}_r$.

We note first that if $\mathfrak{L}$ is any subring of $\mathfrak{A}$ containing $\mathbf{P}_r$, then

[9] [3], p. 4.

$K = \mathfrak{L} \cap G$ is a closed subgroup of G. Evidently K is closed under multiplication. If $A \in K$, A has an inverse. Hence A is not a zero-divisor in $\mathfrak{L}$. Since $\mathfrak{L}$ satisfies the chain conditions this implies that $A^{-1} \in \mathfrak{L}$. Hence K is a group. Now let $I_1 = \tau_{1l}\tau_{1r}^{-1}$ and $I_2 = \tau_{2l}\tau_{2r}^{-1} \in K$. Then $\tau_{3l} = I_1\tau_{1r} + I_2\tau_{2r} \in \mathfrak{L}$ and if $\tau_3 \neq 0$, $I_3 = \tau_{3l}\tau_{3r}^{-1} \in \mathfrak{L}$. Since G is closed $I_3 \in G$. Thus $I_3 \in K$. It follows that K is a closed group of automorphisms.

Next let K be any closed subgroup of G and let $\mathfrak{L} = K\mathbf{P}_r$. Clearly $\mathfrak{L}$ is a subring of $\mathfrak{A}$ containing $\mathbf{P}_r$. Let $K' = \mathfrak{L} \cap G$. Then $K' \supseteq K$. On the other hand the elements of K' leave invariant every K-invariant. Hence, by Theorem 5, $K' \subseteq K$. Thus $K' = \mathfrak{L} \cap G = K$.

Finally let $\mathfrak{L}$ be any ring between $\mathfrak{A}$ and $\mathbf{P}_r$. Let $K = \mathfrak{L} \cap G$ and let $\mathfrak{L}' = K\mathbf{P}_r$. We wish to show that $\mathfrak{L}' = \mathfrak{L}$. Let $L = \mathfrak{L} \cap H$. Then L is the subgroup of inner automorphisms in K. We may choose representatives $A_1, A_2, \cdots, A_u$ of the cosets of H in G such that $A_1, \cdots, A_t$ are representatives of the cosets of L in K. Also we may choose the inner automorphisms $I_1 = \tau_{1l}\tau_{1r}^{-1}, \cdots, I_h = \tau_{hl}\tau_{hr}^{-1}$ in such a way that $(\tau_1, \cdots, \tau_h)$ is a basis for the algebra $\mathbf{A}$ associated with G while $(\tau_1, \cdots, \tau_f)$ is one for the algebra $\mathbf{B}$ associated with K. Now suppose that the element $V = \Sigma A_i I_j \mu_{ji r} \in \mathfrak{L}$. We wish to show that $V \in \mathfrak{L}'$. Since $A_i I_i \in K$ if $i \leqq t$, and $j \leq f$ it suffices to assume that $\mu_{ji} = 0$ for $j \leq f$, $i \leq t$ and to prove under this assumption that if $V \in \mathfrak{L}$, $V = 0$.

Suppose then that there exists a $V \neq 0$ in $\mathfrak{L}$ such that if $V = \Sigma A_i I_j \mu_{ji r}$ then $\mu_{ji} = 0$ for $j \leq f$, $i \leq t$. We may assume also that V is an element satisfying these conditions that has the fewest number of non-zero coefficients $\mu_{ji r}$. Then by the argument used to prove Lemma 2 we see that $V = A_i \sum_j I_j \mu_{ji r} = A_i \Sigma \tau_{jl} \tau_{jr}^{-1} \mu_{ji r}$. By the argument used to prove Lemma 1 we can see that we may assume that the elements $\gamma_j = \tau_j^{-1}\mu_{ji} \in \Gamma$. Hence $V = A_i \Sigma \tau_{jl}\gamma_{jl}$ and since $V \neq 0$, $\tau = \Sigma \tau_j \gamma_j \neq 0$. Thus $V = A_i I \tau_r$ where $I = \tau_l \tau_r^{-1}$. Evidently $A_i I \in K = \mathfrak{L} \cap G$. Hence $i \leq t$ and τ is a linear combination of the τ_j with $j \leq f$. This contradicts the assumption on the form of V. We have therefore proved the following theorem:

THEOREM 7. *Let G be a closed group of automorphisms of finite reduced order and let $\mathfrak{A} = G\mathbf{P}_r$. Then any subring $\mathfrak{L}$ of $\mathfrak{A}$ containing $\mathbf{P}_r$ has the form $K\mathbf{P}_r$ where K is a closed subgroup of G. Moreover, distinct K's give rise in this way to distinct rings $\mathfrak{L}$.*

Let Φ be the division ring of G-invariants. Then we know that $\mathfrak{A}$ is the complete ring of l. t. of $\mathbf{P}$ over Φ_l. We have also seen that there is a $(1-1)$ correspondence between the rings $\mathfrak{L}$ between $\mathfrak{A}$ and $\mathbf{P}_r$ and the division rings

between Φ and $\mathfrak{P}$. If we combine this $(1-1)$ correspondence with the $(1-1)$ correspondence between the rings $\mathfrak{L}$ and the closed subgroups K of G we obtain the following result which includes the basic theorem of the ordinary Galois theory:

THEOREM 8. *Let $\mathfrak{P}$ be a division ring and let G be a closed group of automorphisms in $\mathfrak{P}$ of finite reduced order. To each closed subgroup K of G we associate the division subring $\Psi = \mathfrak{P}(K)$ of elements that are invariant under every $B \in K$ and to each division subring Ψ of $\mathfrak{P}$ containing $\Phi = \mathfrak{P}(G)$ we associate the closed subgroup $K = G(\Psi)$ of automorphisms of G that leave the elements of Ψ invariant. Then these two correspondences are inverses of each other, and each is $(1-1)$ between the closed subgroups of G and the division rings between $\mathfrak{P}$ and Φ.*

7. It is an unsolved problem to determine conditions on a division subring Φ of $\mathfrak{P}$ in order that there exist a closed group G of finite reduced order whose ring of invariants is Φ. If $\mathfrak{P}$ is commutative we know that necessary and sufficient conditions are that $\mathfrak{P}$ be finite, separable and normal over Φ.

Suppose now that $(\mathfrak{P} : \Gamma) < \infty$ and that Φ is a subfield of Γ such that $(\Gamma : \Phi) < \infty$ and let G be a closed group of automorphisms of $\mathfrak{P}$ that has Φ as ring of invariants. Let H be the closed subgroup of G of automorphisms that leave the elements of Γ invariant. Now it is well known that an automorphism that produces the identity effect in Γ is inner. Hence H coincides with the subgroup of inner automorphisms of G. If $\mathfrak{A}$ is the subalgebra associated with H we know that $(\mathfrak{A} : \Gamma) = (\mathfrak{P} : \Gamma)$. Hence $\mathfrak{A} = \mathfrak{P}$. Thus G includes all the inner automorphisms. The automorphisms A of G induce automorphisms $\bar{A}$ in Γ and the correspondence A to $\bar{A}$ is a homomorphism of G on a group $\bar{G}$ of automorphisms in Γ. The kernel of this homomorphism is the set of automorphisms that leave the elements of Γ invariant. Hence the kernel is H and $\bar{G} \cong G/H$. Now Φ is the field of invariants of the group G acting in Γ. Hence Γ is finite, separable and normal over Φ.

Conversely suppose that $\mathfrak{P}$ is finite over Γ and that Γ is finite, separable and normal over Φ. Assume, moreover, that every automorphism $\bar{A}$ of the Galois group of Γ over Φ can be extended to an automorphism A in $\mathfrak{P}$. Let G be the totality of automorphisms of the form AI where I is inner. If $\bar{A}\bar{B} = \bar{C}$ and A, B, C, respectively, are the extensions of $\bar{A}, \bar{B}, \bar{C}$ then AB and C differ by an inner automorphism. Hence $AB = CI$. It follows easily that G is a group of automorphisms containing all of the inner automorphisms and that Φ is the ring of G-invariants.

Thus we have proved the following result which is in essence due to

Teichmüller:[10] If Φ is a subfield of the center Γ of $\mathbf{P}$ and $(\mathbf{P} : \Phi)$ is finite then necessary and sufficient conditions that there exist a closed group G of automorphisms whose set of invariants is Φ are 1) Γ is separable and normal over Φ and 2) every automorphism of the Galois group of Γ over Φ can be extended to an automorphism in $\mathbf{P}$.

When these conditions hold we have seen that G is a group extension of the group H of inner automorphisms of $\mathbf{P}$ by the finite Galois group $\bar{G}$ of Γ over Φ. It is natural to seek conditions that this extension split in the sense that G contains a subgroup G' whose elements are representatives of the cosets of $G/H \cong \bar{G}$. It is easy to see that this is equivalent to the requirement that $\mathbf{P}$ be a direct product $\Sigma \times \Gamma$ where Σ and Γ are algebras over Φ. Conditions for this have been given by Teichmüller.[11].

THE JOHNS HOPKINS UNIVERSITY.

BIBLIOGRAPHY.

[1] E. Artin and G. Whaples, "The theory of simple rings," *American Journal of Mathematics*, vol. 65 (1943), pp. 87-107.

[2] N. Jacobson, *The Theory of Rings*, Mathematical Surverys, vol. 2, New York, 1943.

[3] ———, "The fundamental theorem of the Galois Theory for quasi-fields," *Annals of Mathematics*, vol. 41 (1940), pp. 1-7.

[4] ———, "Structure theory of simple rings without finiteness assumptions," *Transactions of the American Mathematical Society*, vol. 57 (1945), pp. 228-245.

[5] E. Noether, "Nichtkommutative Algebra," *Mathematische Zeitschrift*, vol. 37 (1933), pp. 514-541.

[6] O. Teichmüller, "Über die sogenannte nichtkommutative Galoissche Theorie etc.," *Deutsche Mathematik*, vol. 5 (1940), 138-149.
See also the review in the *Zentralblatt*, vol. 23, p. 198.

---

[10] [6].

[11] [6].

THE DOUBLE CHAIN CONDITION IN CYCLIC OPERATOR GROUPS.*

By REINHOLD BAER.

In attempting to generalize Fitting's Lemma [1] the author found [2] that this was possible for abelian operator groups meeting the following requirements:

(a) the descending chain condition is satisfied by the admissible subgroups;

(b) if Z is a cyclic subgroup (i. e., if Z is generated by one element), then the ascending chain condition is satisfied by the admissible subgroups of Z.

Naturally one wonders whether condition (b) is a consequence of (a); and added strength is given to such a conjecture, if one remembers Hopkins' [3] Theorem to the effect that in rings possessing an identity the ascending chain condition for right-ideals is a consequence of the descending chain condition. However, it is possible to construct examples of cyclic groups where the descending, though not the ascending, chain condition is satisfied by the admissible subgroups.

In the light of Hopkins' Theorem, just quoted, it is only natural to assume that there will exist a large class of groups where cyclicity and descending chain condition imply the ascending chain condition; and it is the object of the present note to exhibit such classes of groups.

Let A be an abelian group where the composition is written as addition $a + b$; and assume that A admits the elements in the system M as operators (= right multipliers). The M-subgroup S of A is said to be of *length* $n = n(S)$, if the M-group S possesses a composition series [4] of length n.

An M-subgroup J of A is termed *minimal*, if $J \neq 0$ and if J does not

* Received January 3, 1946. Presented to the American Mathematical Society December 26, 1946.

[1] See Jacobson (1), p. 9 for a statement of Fitting's Lemma. Various applications of this proposition may be found in Jacobson's book.

[2] Baer (2).

[3] Apart from Hopkins' original paper, see Baer (1) for a proof, under weakened hypotheses, or Jacobson (1), p. 71, Theorem 29.

[4] See e. g., Jacobson (1), p. 7.

possess any M-subgroups apart from 0 and J. This is equivalent to saying that minimal M-subgroups are the M-subgroups of length 1. The sum of all the minimal M-subgroups of A will be called the *socle* [5] $S(A)$ of A. It is well known [6] that double chain condition, ascending chain condition and descending chain condition for admissible subgroups are equivalent properties of the socle.

The ascending Loewy series $L(A,i)$ of A may now be defined [7] inductively as follows: $L(A,0)=0, L(A,i+1)$ is the uniquely determined M-subgroup of A satisfying: $L(A,i+1)/L(A,i)=S(A/L(A,i))$. The subgroups $L(A,i)$ form an ascending chain of M-subgroups of A so that their (set-theoretical) join is an M-subgroup $L(A,\omega)$ of A. It is clear how to define the Loewy series for tranfinite ordinals as well; but we have no need to do so.

LEMMA 1. *$L(A,i)$ contains every M-subgroup of A whose length does not exceed i.*

Proof. 0 is the only subgroup of length 0, proving our contention for $i=0$. Thus we may assume the validity of our contention for $i-1$ in order to prove it for i. If T is some M-subgroup of A such that $n(T)\leqq i$, then T possesses a maximal M-subgroup T' so that $n(T)=n(T')+1$. Since, therefore, $n(T')\leq i-1$, it follows from our induction hypothesis that $T'\leq L(A,i-1)$. Since T' is part of T also, this implies that $T'\leq T\cap L(A,i-1)\leq T$. But T/T' is of length 1; and thus $T/[T\cap L(A,i-1)]$ is of length 0 or 1. This latter group is isomorphic to $[T+L(A,i-1)]/L(A,i-1)$ so that this last group is either 0 or a minimal subgroup of $A/L(A,i-1)$. It is therefore part of the socle of $A/L(A,i-1)$ proving that $T+L(A,i-1)\leq L(A,i)$. Hence $T\leq L(A,i)$ whenever $n(T)\leq i$, completing the proof.

COROLLARY 1. *If the double chain condition is satisfied by the M-subgroups of $L(A,i)/L(A,i-1)$ for every positive i, then the following properties of the M-group A imply each other:*

(i) $A=L(A,\omega)$.

(ii) *The cyclic M-subgroups* [8] *of A are of finite length.*

Proof. It is readily seen that every $L(A,i)$ is of finite length. If (i)

---

[5] Following Remak. Note that the *socle* may be equal to 0.

[6] See e. g. Jacobson (1), p. 14, Theorem 12.

[7] Following Krull.

[8] An M-group is termed cyclic, if it is generated, as an M-group, by one element.

is satisfied, then every cyclic M-subgroup C of A is part of some $L(A,i)$. But $L(A,i)$ if of finite length and thus every M-subgroup of $L(A,i)$ is of finite length. Hence C is of finite length, proving that (ii) is a consequence of (i).

If, conversely, (ii) is valid, then it follows from Lemma 1 that every element in A belongs to some $L(A,i)$; and thus (i) is a consequence of (ii).

COROLLARY 2. *The double chain condition is satisfied by the M-subgroups of A if, and only if,*

(a) *the double chain condition is satisfied by the M-subgroups of $L(A,i)/L(A,i-1)$ for every positive i; and*

(b) *there exists an integer N such that every cyclic M-subgroup of A is of length not exceeding N.*

Proof. If the double chain condition is satisfied by the M-subgroups of A, then A has a definite finite length N which is an upper bound for the length of all its M-subgroups. This proves the necessity of condition (b); and the necessity of (a) is a consequence of the well known fact that subgroups and quotient groups of groups of finite length are of finite length too.

If, conversely, conditions (a) and (b) are satisfied, then we infer from Lemma 1 that $A = L(A,N)$ by (b); and this group is of finite length by (a).

Example 1. Denote by p a prime number and by A a direct sum of a countable infinity of cyclic groups of order p. Then A possesses a basis $b(1), b(2), \cdots, b(i), b(i+1), \cdots$. There exists an endomorphism $s(i)$ of A such that $b(j)s(i) = 0$ for $j \neq 2i+1$ and $b(2i+1)s(i) = b(2i)$; and there exists an endomorphism $t(i)$ of A such that $b(j)t(i) = 0$ for $j \neq 2i+2$ and $b(2i+2)t(i) = b(2i)$. The $s(i)$ and $t(i)$ for $0 < i$ form the system M.

It is readily seen that $L(A,i)$ for $0 < i$ is generated by the elements $b(1), b(3), \cdots, b(2i-1), b(2i)$; and that $b(2i-1), b(2i)$ represent a basis of $L(A,i)/L(A,i-1)$ so that this group is of length 2. Thus the M-group A meets the requirement (a) of Corollary 2 in the strengthened form:

(a′) *the groups $L(A,i)/L(A,i-1)$ for $0 < i$ are of bounded length.*

Since every element in A belong to some $L(A,i)$, we see that the following form of condition (b) of Corollary 2 is satisfied by A:

(b′) $A = L(A,\omega)$.

But it is clear that the ascending chain condition is not satisfied by the M-subgroups of A, since $L(A,i-1) < L(A,i)$ for $0 < i$. The descending chain condition is not satisfied by the M-subgroups of A either, since $b(2i-1)$ for $0 < i$ is not contained in the M-subgroup of A which is generated by the elements $b(2i), b(2i+1), \cdots, b(2i+j), \cdots$.

Note, furthermore, that (b′) is just the condition (i) of Corollary 1 and that (a′) implies the general hypothesis of Corollary 1. Thus the groups meeting all the requirements of Corollary 1 need not satisfy any of the chain conditions.

Following H. Prüfer we say that *the M-group A is of finite rank, if there exists an integer R with the following property:*

(R) *If U is a finite subset of A, then there exists an M-subgroup V of A which is generated by not more than R elements and contains U.*

Every finitely generated M-group is of finite rank through the converse is not true, as may be seen from the example of the groups of type p^∞. Quotient groups of groups of finite rank are of finite rank too; but subgroups of groups of finite rank need not be a finite rank, as may be seen from the following

Example 2. If p is a prime number, A a direct sum of a countable infinity of cyclic groups of order p, and if $b(0), b(1), \cdots, b(i), \cdots$ is a basis of A, then there exists one, and only one, endomorphism $a(i)$ for $0 < i$ which maps $b(0)$ upon $b(i)$ and $b(j)$ for $0 < j$ upon 0. Denote by M the system consisting of the endomorphisms $a(i)$.

Clearly A is a cyclic M-group generated by $b(0)$. Denote by B the subgroup of A which is generated by the elements $b(i)$ for $0 < i$. Every subgroup of B is M-admissible; and it is clear that B is not of finite rank.

This Example 2 shows furthermore that we would have obtained a narrower class of groups, if we had substituted for condition (R) the following stricter condition.

(R′) If the M-subgroup U of A is finitely generated, then it may be generated by R of its elements.

LEMMA 2. *The double chain condition is satisfied by the M-subgroups of A if, and only if,*

(a) *A is of finite rank; and*

(b) *the lengths of the cyclic M-subgroups of A are bounded.*

Proof. If the double chain condition is satisfied by the M-subgroups of A, then A is finitely generated, proving the necessity of (a); and the necessity of (b) is a consequence of Corollary 2.

Suppose, conversely, the validity of conditions (a) and (b). Denote by R an integer such that every finite subset of A is contained in a M-subgroup of A which is generated by not more than R elements; and denote by N an integer such that the length of every cyclic M-subgroup of A is at most N. If an M-subgroup of A is generated by not more than i elements, then it is clearly of a length not exceeding iN; and the length of every finitely generated M-subgroup of A consequently does not exceed RN. Hence there exists an M-subgroup B of greatest length (note that M-subgroups of finite length are always finitely generated). If Z is a cyclic M-subgroup of A, then Z is of finite length and hence $B+Z$ is of finite length. This length cannot be greater than the length of B, proving (by the Jordan-Hölder Theorem) that $Z \leq B$. Hence $B = A$ so that A itself is of finite length, as we wished to show.

Again it is easy to show that neither of the conditions (a) and (b) can be omitted.

It is clear on reading the proofs of Lemmas 1 and 2 that the arguments used are purely lattice theoretical. It would be easy, though too lengthy for our purposes, to state these lemmas as theorems on lattices.

Example 3. Denote by p a prime number and let A be a direct sum of a countable infinity of cyclic groups of order p. Denote by $b, b(1), b(2), \cdots, b(i), \cdots$, for $0 < i$, a basis of A. Then there exists one and only one endomorphism $a(i)$ of A which maps b upon $b(i)$ and which maps every $b(j)$ upon 0; and there exists one and only one endomorphism $\alpha(i)$ of A which maps $b(i+1)$ upon $b(i)$ and which maps b as well as $b(j)$, for $j \neq i+1$, upon 0. Denote by M the system consisting of the $a(i)$ and $\alpha(i)$.

It is clear that A is a cyclic M-group, generated by b. Denote by $S(i)$ the subgroup of A which is generated by $b(1), \cdots, b(i)$; and by $S(\omega)$ the subgroup of A which is generated by all the $b(i)$. These subgroups of A are readily seen to be M-admissible; and they satisfy:

$$0 < S(1) < \cdots < S(i) < S(i+1) < \cdots < S(\omega) < A.$$

Consider now some M-subgroup S of A. If S is not part of $S(\omega)$, then S contains an element of the form: $b + b'$ where b' is in $S(\omega)$. But then S contains $(b+b')a(i) = b(i)$ for every i. Hence b' and therefore b are in S, proving $S = A$, if S is not part of $S(\omega)$.

If the M-subgroup S of A contains an element that belongs to $S(i+1)$, but not to $S(i)$, then S contains an element of the form $b(i+1)+z$ for z in $S(i)$. Consequently S contains $(b(i+1)+z)\alpha(i) = b(i), b(i)\alpha(i-1) = b(i-1)$ etc., so that $S(i) \leqq S$. Hence z is in S also which implies that $b(i+1)$ belongs to S. Consequently $S(i+1) \leqq S$.

From the results of the last two paragraphs we deduce that 0, $S(i)$, $S(\omega)$ and A are the only M-subgroups of A. Thus we see that A is a cyclic M-group and that the descending, but not the ascending chain condition is satisfied by the M-subgroups of A.

If the abelian group A admits the elements in the system M as operators, then denote by M^* the ring of endomorphisms of the ordinary abelian group A which is generated by the identity and by those endomorphisms of A which are induced in A by elements in M. A subgroup of A is M-admissible if, and only if, it is M^*-admissible. Hence it does not involve any loss in generality, if we assume in the future that M itself is *a ring containing an identity* 1 *which acts as an identity on* A so that $x1 = x$ for x in A. Nothing will be gained, however, by assuming that the elements in M are endomorphisms of A, since endomorphism rings of A are, in general, not endomorphism rings of the M-subgroups of A. [For different endomorphisms of A may induce the same endomorphism in a subgroup.]

If Z is part of M, then it is possible to consider the M-group A at the same time as a Z-group; and it will be possible to gain information about the M-group A from information concerning the Z-group A, since every M-subgroup of A is at the same time a Z-group of A (though not conversely). From the remarks in the preceding paragraph it follows that we may assume, without loss of generality, that *Z is a subring of M which contains the identity of the ring M*; and we shall make this hypothesis throughout.

If Z is a subring of M and if x is an element in the M-group A, then we denote by $Q(x, Z)$ the set of all the elements z in Z, satisfying $xz = 0$. Clearly $Q(x, Z)$ is a right-ideal in Z which may be called the *Z-order of x*.

LEMMA 3. *If $A = gM$ is a cyclic M-group and if Z is a subring of M such that*

(a) $MQ(g, Z) \leqq Q(g, Z)M$,

(b) *the descending chain condition is satisfied by the Z-subgroups of gZ; then the cyclic Z-subgroups of A are of bounded length.*

Proof. From the definition of the Z-order and from condition (a) we infer that

$$gZ\,Q(g, Z) \leqq gMQ(g, Z) \leqq g\,Q(g, Z)M = 0.$$

Hence $ZQ(g,Z) \leq Q(g,Z)$ so that $Q(g,Z)$ is a two-sided ideal in Z. Mapping the element gy, for y in Z, upon the coset $Q(g,Z)+y$ we obtain an isomorphism of the Z-group gZ upon the Z-group $Z/Q(g,Z)$; and this isomorphism induces, in particular, an isomorphism of the lattice of Z-subgroups of gZ upon the lattice of right-ideals in the ring $Z/Q(g,Z)$. It follows from (b) that the descending chain condition is satisfied by the right-ideals of the ring $Z/Q(g,Z)$, a ring with unit. Applying Hopkins' Theorem [9] we see now that the double chain condition is satisfied by the right-ideals in $Z/Q(g,Z)$. Using the isomorphism between gZ and $Z/Q(g,Z)$ we conclude that

the cyclic Z-group gZ is of finite length N.

If x is an element in $A=gM$, then there exists an element m in M such that $x=gm$. Hence it follows from (a) that

$$xQ(g,Z)=gmQ(g,Z)\leqq gMQ(g,Z)\leqq gQ(g,Z)M=0.$$

We deduce, therefore, from the definition of the Z-order that

$$Q(g,Z)\leqq Q(x,Z) \text{ for every } x \text{ in } A.$$

As before we see that the Z-groups xZ and $Z/Q(x,Z)$ are essentially the same; and that, therefore, the lattices of Z-subgroups of xZ is isomorphic to the lattice of right-ideals between $Q(x,Z)$ and Z. From $Q(g,Z)\leqq Q(x,Z)$ and the fact that the lattice of right-ideals between $Q(g,Z)$ and Z is of length N, we infer now that the lattices of right-ideals between $Q(x,Z)$ and Z is of a length not exceeding N. The Z-group xZ is, therefore, of length not greater than N; and this completes the proof.

THEOREM 1. *The double chain condition is satisfied by the M-subgroups of the cyclic M-group $A=gM$, if there exists a subring Z of M with the following properties:*

(i) $MQ(g,Z)\leqq Q(g,Z)M$,

(ii) *the descending chain condition is satisfied by the Z-subgroups of A.*

Proof. It is an immediate consequence of conditions (i) and (ii) and of Lemma 3 that the conditions (a) and (b) of Corollary 2 are satisfied by the Z-subgroups of A. It follows from Corollary 2 that the double chain condition is satisfied by the Z-subgroups of A. But the system of M-subgroups

[9] Which we stated in the introduction.

of A is part of the system of Z-subgroups of A, proving the validity of the double chain condition for the M-subgroups of A.

COROLLARY 3. *The double chain condition is satified by the M-subgroups of the cyclic M-group A, if either of the following conditions is satisfied by A and M:*

(1) *Every right-ideal in M is two-sided; and the descending chain condition is satisfied by the M-subgroups of A.*

(2) *If C is the center of M, then the descending chain condition is satisfied by the C-subgroups of A.*

This is an immediate consequence of Theorem 1.

The ring M shall be said to be *of finite rank over its subring Z*, if the addition group M_+ of M is a Z-group [10] of finite rank.

LEMMA 4. *If $A = gM$ is a cyclic M-group, and if M is of finite rank over its subring Z, then A is a Z-group of finite rank.*

Proof. There exists an integer R with the following property:

If F is a finite subset of M, then there exist elements $m(1), \cdots, m(k)$ in M such that $k \leq R$ and such that $F \leq \sum_{i=1}^{k} m(i)Z$.

Consider now a finite subset T of A. Then there exists a finite subset F of M such that $T = gF$. Hence there exist elements $m(1), , \cdots, m(k)$ such that $k \leq R$ and $F \leq \sum_{i=1}^{k} m(i)Z$. Clearly $T = gF \leq \sum_{i=1}^{k} [gm(i)]Z$, proving that the Z-group A is of finite rank.

THEOREM 2. *The double chain condition is satisfied by the M-subgroups of the cyclic M-group $A = gM$, if there exists a subring Z of M with the following properties:*

(i) $MQ(g,Z) \leqq Q(g,Z)M$;

(ii) *the descending chain condition is satisfied by the Z-subgroups of gZ;*

(iii) *M is of finite rank over Z.*

Proof. It is a consequence of Lemmas 3 and 4 that conditions (a) and (b) of Lemma 2 are satisfied by the Z-subgroups of A. We infer from

[10] The elements in Z act as right multipliers only for M_+.

Lemma 2 the validity of the double chain condition for the Z-subgroups of A, implying the double chain condition for M-subgroups.

COROLLARY 4. *The double chain condition is satisfied by the M-subgroups of the cyclic M-group $A = gM$, if M is of finite rank over its center C and if the descending chain condition is satisfied by the C-subgroups of gC.*

This is an immediate consequence of Theorem 2.

UNIVERSITY OF ILLINOIS,
URBANA, ILLIOIS.

BIBLIOGRAPHY.

R. Baer.

(1) "Radical Ideals," *American Journal of Mathematics*, vol. 64 (1943), pp. 537-568.

(2) "Splitting Endomorphisms," *Transactions of the American Mathematical Society.*

Ch. Hopkins.

(1) "Rings with minimal condition for left ideals," *Annals of Mathematics*, vol. 40 (1939), pp. 712-730.

N. Jacobson.

(1) *The theory of rings*, Mathematical Surveys, No. 2, 1943.

RADICALS AND SUBDIRECT SUMS.*[1]

By BAILEY BROWN and NEAL H. MCCOY.

1. Introduction. The radical N of a ring R, in which the right (left) ideals satisfy the descending chain condition, is usually defined as the join of all nil right (left) ideals in R. It is well known that N is a two-sided ideal in R and that the ring R/N is semi-simple, that is, has zero radical. Furthermore, a necessary and sufficient condition that R be semi-simple is that R be isomorphic to a full direct sum of a finite number of simple rings, each of which is a complete matrix ring over a division ring (Wedderburn-Artin theorems).

Various definitions have been proposed for the radical of an arbitrary ring,[2] most of which are in terms of nil ideals. Jacobson, however, has abandoned nilpotence as fundamental in the notion of radical;[3] instead, his definition is in terms of a concept first introduced by Perlis.[4] If N' is the Jacobson radical of R and $N' \neq R$, that is, R is not a radical ring, it turns out that N' is the intersection of a set of two-sided ideals B in R such that R/B is of a special type called primitive. From this, it follows that $N' = 0$ if and only if R is isomorphic to a subdirect sum of primitive rings. In the presence of the descending chain condition for right ideals, a primitive ring is necessarily simple, and a subdirect sum of a finite number of simple rings is actually a full direct sum of a finite number of these rings. Jacobson's results thus generalize the Wedderburn-Artin structure theorems mentioned above.

Our point of view, based somewhat on the form of Jacobson's results, is that the radical N of an arbitrary ring should be a two-sided ideal whose vanishing is a necessary and sufficient condition that R be isomorphic to a subdirect sum of rings of some particular type. Furthermore, if the descending chain condition holds for right ideals in R the radical should coincide with the classical one. We shall presently define a radical which generally differs from Jacobson's but which, in common with his, meets both these conditions.

* Received May 18, 1946.
[1] Presented to the Society, April 27, 1946.
[2] Köthe [1], Levitzki [1] and [2], Baer [1].
[3] Jacobson [1].
[4] Perlis [1].

First, however, we state his definition in terms which will conveniently lead to ours.

Let a be an arbitrary element of R, and consider the right ideal $\{ax - x\}$, where x runs over R.[5] Now b is in the radical N' of R, as defined by Jacobson, if and only if $a \in \{ax - x\}$ for *every* element a in the right ideal generated by b. We shall henceforth refer to N' as the Jacobson radical of R.

Now let $G(a)$ be the two-sided ideal in R generated by the right ideal $\{ax - x\}$, that is,

$$G(a) = \{ax - x + \Sigma y_i a z_i - \Sigma y_i z_i\},$$

where x, y_i, z_i vary over R, the sums naturally being finite. We shall say that b is an element of the *radical* N of R if and only if $a \in G(a)$ for every element a of the two-sided ideal generated by b. Unless otherwise stated, this is henceforth the meaning we attach to the word "radical." We show that N is a two-sided ideal in R, that R/N has zero radical, and that the radical of a complete matrix ring R_n is N_n. *The vanishing of N is a necessary and sufficient condition that R be isomorphic to a subdirect sum of simple rings with unit element.*

If $b \in N'$, and a is an arbitrary element of the two-sided ideal generated by b, then, since N' is itself a two-sided ideal, it follows that $a \in \{ax - x\}$, so that $a \in G(a)$ and $b \in N$. Thus $N' \subseteq N$ and, since a primitive ring need not be a simple ring with unit element, in general, N' is properly contained in N. However, the definitions coincide if R is commutative, and hence for commutative rings, $N' = N$. We shall show that also $N' = N$ if the descending chain condition holds for right ideals in R. Thus our results furnish another generalization of one of the Wedderburn-Artin structure theorems and also, in common with Jacobson's, yield as special cases a number of known results on subdirect sums.[6]

Actually, we use a somewhat more general approach which may be of some interest in itself. In the proofs of some of the theorems about the radical, the only property of the $G(a)$ defined above which is actually used is that if $a \to \bar{a}$ is a homomorphism of R onto $\bar{R}$, then $G(\bar{a}) = \overline{G(a)}$. Accordingly, we consider an arbitrary mapping $a \to F(a)$, defined for all rings, of R into the set of two-sided ideals in R such that $F(\bar{a}) = \overline{F(a)}$ for every homomorphism $a \to \bar{a}$ of R onto a ring $\bar{R}$. If, then, in the definition of radical, we use $F(a)$ in place of $G(a)$, we get what we shall call the *F-radical* of R; thus the radical

[5] Jacobson uses $\{ax + x\}$, but we find it convenient to use this equivalent formulation.

[6] Stone [1], McCoy [1], McCoy and Montgomery [1].

is the G-radical. It is shown that the F-radical N_F of R is a two-sided ideal in R and that R/N_F has zero F-radical. *The vanishing of N_F is a necessary and sufficient condition that R be isomorphic to a subdirect sum of subdirectly irreducible rings of zero F-radical.*[7] Although the F-radical has a number of "radical-like" properties, it will not in general reduce to the ordinary radical in the presence of the descending chain condition. The concept is, therefore, not a true generalization of the radical, but rather a device leading to subdirect sum decompositions of various types, depending on the choice of the mapping F.

In view of these results, it becomes of some interest to characterize the subdirectly irreducible rings of zero F-radical. In the final section, we do this for several different mappings $a \to F(a)$, thus obtaining some other interesting special cases of the general subdirect sum decomposition theorem stated above.

Inasmuch as we shall be chiefly concerned with two-sided ideals the word *ideal* shall henceforth, unless otherwise stated, mean *two-sided ideal.*

2. Subdirect sum of rings. An element of the (full) direct sum S of rings R_i will be denoted by

$$a = (a_1, a_2, \cdots),$$

where $a_i \in R_i$ and i, despite the notation employed, runs over an index set of arbitrary cardinal number. By a *subdirect sum* of the rings R_i is meant a subring T of S such that, for each i, the homomorphism $a \to a_i$ of S onto R_i maps T on all of R_i, inducing a homomorphism of T *onto* R_i. If a ring R is isomorphic to T, the product of the isomorphism and this homomorphism is a *natural* homomorphism of R onto R_i. *A ring R is isomorphic to a subdirect sum of rings R_i if and only if, for each i, R contains an ideal M_i with $R/M_i \cong R_i$, and the ideals M_i have zero intersection.*[8]

By a *subdirectly irreducible* ring[7] is meant a ring R such that in every isomorphic representation of R as a subdirect sum of rings R_i, the natural homomorphism of R onto R_i is an isomorphism for some i. Thus, *R is subdirectly irreducible if and only if the intersection of all nonzero ideals in R is itself a nonzero ideal J.* This ideal J is a unique minimal ideal in R if R has more than one element and is vacuous otherwise.

---

[7] The concept of subdirectly irreducible ring was introduced by Birkhoff [1]. See **2** below.

[8] McCoy and Montgomery [1], McCoy [1].

3. The F-radical of a ring. Suppose that in each ring R there is assigned a mapping F of R into the set of ideals in R in such a way that if $a \to \bar{a}$ defines a homomorphism of R onto a ring $\bar{R}$, and $\overline{F(a)}$ is the image in $\bar{R}$ of the ideal $F(a)$ in R, then $F(\bar{a}) = \overline{F(a)}$. An element a of R shall be *F-regular* if and only if $a \in F(a)$, and an ideal in R shall be *F-regular* if and only if each of its elements is F-regular.

DEFINITION. The *F-radical* N_F of R is the set of elements b of R which generate F-regular ideals (b) in R.

If $N_F = R$, R is an *F-radical ring.*

If $a \in F(a)$, then $\bar{a} \in \overline{F(a)} = F(\bar{a})$. It follows that F-regularity of elements, and therefore of ideals, is preserved under homomorphism, and that the F-radical of R goes *into* the F-radical of $\bar{R}$.

The following theorem characterizes the F-radical:[9]

THEOREM 1. *The F-radical N_F of a ring R is the intersection ΠM of the ideals M in R such that R/M is subdirectly irreducible and has zero F-radical.*

We show first that $\Pi M \subseteq N_F$. If $b \notin N_F$, then for some a in (b), $a \notin F(a)$. By Zorn's Maximum Principle, there is an ideal M containing $F(a)$ but not containing a, such that every ideal in R which contains M as proper subset does contain a. If $\bar{a}$ is the residue class to which a belongs modulo M, it follows that every nonzero ideal in R/M contains $\bar{a} \neq 0$, so R/M is subdirectly irreducible. But since $F(\bar{a}) = \overline{F(a)} = \bar{0}$, $\bar{a} \notin F(\bar{a})$. Thus every nonzero ideal in R/M contains an element which is not F-regular, so that R/M has zero F-radical. Since $a \notin M$, it follows that $b \notin M$, and $b \notin \Pi M$.

Conversely, $N_F \subseteq \Pi M$. For if $b \in N_F$ and M is any ideal of the specified type, then $\bar{b}$ is in the F-radical of R/M since F-regularity is preserved under homomorphism. Since R/M has zero F-radical, $\bar{b} = \bar{0}$, $b \in M$. Thus $b \in \Pi M$.

COROLLARY 1. *The F-radical of R is an ideal in R.*

COROLLARY 2. *A ring R is an F-radical ring if and only if R itself is the only ideal M in R such that R/M is subdirectly irreducible of zero F-radical.*

COROLLARY 3. *The F-radical of R is the join of all F-regular ideals in R.*

This last statement follows from the fact that, by definition, N_F consists of the elements of R which generate F-regular ideals. Since N_F is an ideal it is clearly an F-regular ideal which contains every F-regular ideal.

[9] Cf. Birkhoff [1]; Jacobson [1], p. 311.

THEOREM 2. *The ring R/N_F has zero F-radical.*

Let $\bar{b}$ denote an element of the F-radical of R/N_F, and M any ideal in R such that R/M is subdirectly irreducible and has zero F-radical. Since $R/M \cong (R/N_F)/(M/N_F)$, it follows from Theorem 1 that $\bar{b} \in M/N_F$. Hence $b \in M$, $b \in \Pi M = N_F$, $\bar{b} = \bar{0}$.

THEOREM 3. *A subdirect sum of rings R_i of zero F-radical has zero F-radical.*

If a is in the F-radical of R and $a \to a_i$ is the natural homomorphism of R onto R_i, then, for each i, a_i is in the F-radical of R_i. Hence $a_i = 0$ and $a = 0$.

Almost immediate is

THEOREM 4. *A ring has zero F-radical if and only if it is isomorphic to a subdirect sum of subdirectly irreducible rings of zero F-radical.*

The necessity of the condition follows from Theorem 1 and the results of **2**; the sufficiency, from Theorem 3.

If, for example, the mapping F is defined in each ring R by $F(a) = 0$ for every a in R, zero is the only F-regular element of R and every ring has zero F-radical. Theorem 4 then reduces to the ring case of Birkhoff's theorem on algebras with finitary operations.[7]

We now prove

THEOREM 5. *A subdirectly irreducible ring R has zero F-radical if and only if the minimal ideal J contains an element $a \neq 0$ such that $F(a) = 0$.*

If R has zero F-radical and $j \neq 0$ is an element of J, some element a of $(j) = J$ is not F-regular, that is, $a \notin F(a)$. This element a can not be zero and is contained in every nonzero ideal in R. Hence we must have $F(a) = 0$. Conversely, the F-radical of R does not contain J, for, by hypothesis, J contains an element a which is not F-regular. Hence the F-radical is the zero ideal.

In the next two sections, we make a detailed study of the radical, that is, the case in which the general mapping $a \to F(a)$ is specialized to $a \to G(a)$, where $G(a)$ is as defined in the introduction. In **6**, further illustrations and applications of this general theory will be given.

4. The radical of a ring.

We now repeat the

DEFINITION. The *radical* N of a ring R is the set of all elements b of R such that for every element a of (b),

$$a \in G(a) = \{ax - x + \Sigma y_i a z_i - \Sigma y_i z_i\}.$$

A ring with zero radical will be said to be *semi-simple*; a ring which coincides with its radical is a *radical ring*. If a ring has more than one element and a unit element 1, it is clear that $1 \notin G(1) = 0$, so that such a ring can not be a radical ring.

THEOREM 6. *A subdirectly irreducible ring is semi-simple if and only if it is a simple ring with unit element.*

The theorem is trivial for one-element rings.[10] If the simple ring S has more than one element and a unit element 1, then $1 \in J$, $G(1) = 0$, and Theorem 5 shows that S is semi-simple. Conversely, let S be a subdirectly irreducible ring of more than one element. If S is semi-simple, Theorem 5 states that its minimal ideal J contains an element $e \neq 0$ for which $G(e) = 0$. Hence $ex = x$ for all x in S and $S \subseteq J$, that is, S is simple and has a left unit e. The proof is completed by establishing the following lemma:[11]

LEMMA. *If e is a left unit of a simple ring S, then e is the unit element of S.*

It is given that $ex = x$ for all x in S. The set $E = \{xe - x; x \in S\}$ is a two-sided ideal in S. If $E = S$, then $e = ye - y$ for some y in S, and right multiplication by e yields $e = ye - ye = 0$. Hence $x = ex = 0$ for all x, and S is a one-element ring. If S has more than one element, then $E = 0$, and $x = ex = xe$ for all x in S.

COROLLARY. *A simple ring is semi-simple if it has a unit element, otherwise it is a radical ring.*

The following characterization of the radical follows at once from Theorems 1 and 6:

THEOREM 7. *The radical of a ring R is the intersection of all ideals M in R such that R/M is a simple ring with unit element.*

COROLLARY 1. *If R is not a radical ring, the radical of R is the intersection of all maximal ideals M in R such that R/M has a unit element.*

---

[10] In order to avoid any possible confusion, we explicitly point out that, according to our usage, a one-element ring is a simple ring with unit element.

[11] This was pointed out to us by R. E. Johnson.

COROLLARY 2. *If R is not a one-element ring and has a unit element, the radical of R is the intersection of all maximal ideals in R.*

Theorems 4 and 6 yield at once the following important result:

THEOREM 8. *A ring R is semi-simple if and only if it is isomorphic to a subdirect sum of simple rings with unit element.*

This is, for general rings, a generalization of one of the Wedderburn-Artin theorems for semi-simple rings in which the right ideals satisfy the descending chain condition. We shall show that it actually yields that theorem if merely the descending chain condition for two-sided ideals is assumed.

THEOREM 9. *In the presence of the descending chain condition for two-sided ideals, a semi-simple ring R is isomorphic to the full direct sum of a finite number of simple rings with unit element.*

Since R has zero radical, there exists a set of ideals M in R such that R/M is simple with unit element, and $\Pi M = 0$. Since the descending chain condition is assumed, there is a finite subset $\{M_i\}$, $(i = 1, 2, \cdots, k)$, of these ideals with $\Pi M_i = 0$, and which, furthermore, is minimal in the sense that the intersection of any $k - 1$ ideals of the subset is not zero. If $a \in R$ and a_i is the image in R/M_i of the element a under the natural homomorphism of R onto R/M_i, it is clear that the correspondence

$$a \rightarrow (a_1, a_2, \cdots, a_k) \tag{1}$$

defines an isomorphism of R onto a subdirect sum of the simple rings R/M_i. However, it is easy to show as follows that this subdirect sum is actually the full direct sum. From the minimality of $\{M_i\}$ it is clear that there is an element c of R in $\prod_{i=2}^{k} M_i$, but not in M_1. For this c, the correspondence (1) yields $c \rightarrow (c_1, 0, \cdots, 0)$. Since c_1 is a nonzero element of the simple ring R/M_1, it follows that $(c_1) = R/M_1$. Hence, under the isomorphism (1), every element of the form $(r_1, 0, \cdots, 0)$, $r_1 \in R/M_1$, appears as the image of some element of R. A similar result holds if M_1 is replaced by M_j, $(j = 2, \cdots, k)$. Thus, finally, every element of the direct sum of the rings R/M_i appears as the image of an element of R under the isomorphism (1). This shows that R is isomorphic to the direct sum of the rings R/M_i, $(i = 1, 2, \cdots, k)$.

It was pointed out in the introduction that the Jacobson radical N' is contained in the radical N. We now prove

THEOREM 10. *If the right ideals of R satisfy the descending chain condition, the radical N of R coincides with the Jacobson radical N' of R.*

In view of the remark just made, it is enough to prove that $N \subseteq N'$. In R/N', the descending chain condition holds for right ideals and the Jacobson radical is zero. This implies [3] that R/N' is isomorphic to the direct sum of a finite number of simple rings, each of which has a unit element because of the chain condition. Thus R/N' has zero radical by Theorem 8. Hence if $b \in N$, then $\bar{b}$ is in the radical of R/N', so $\bar{b} = \bar{0}$ and $b \in N'$. Thus $N \subseteq N'$, as required.

If R has more than one element and a unit element, Jacobson has shown that N' is the intersection of all maximal right (left) ideals in R. The second Corollary to Theorem 7, together with the result just established, shows that if R has a unit element and the descending chain condition holds for right ideals in R, the intersection of all maximal right (left) ideals coincides with the intersection of all maximal two-sided ideals.

We shall conclude this section with some further miscellaneous results. First, we remark that if a is a nilpotent element of a ring R, then $a \in G(a)$. For if $a^n = 0$, $n > 1$, and $x = -\sum_{i=1}^{n-1} a^i$, it is easily verified [3] that $a = ax - x$. This establishes

THEOREM 11. *If all elements of (b) are nilpotent, then b is in the radical of R.*

We now pass to the proof of

THEOREM 12. *If A is an ideal in R, the radical of the ring A is contained in the radical N of R.*

Let M be any ideal in R for which R/M is a simple ring with unit element. Then $A/(A \cap M) \cong (A, M)/M$, which, as an ideal in the ring R/M, is itself a simple ring with unit element. Thus, by Theorem 7, the radical of the ring A is contained in $A \cap M$, and therefore in $N = \Pi M$.

COROLLARY. *Any ideal in a semi-simple ring is semi-simple.*

5. The radical of a matrix ring. The ring of all matrices of order n with elements in a ring R will be denoted, as usual, by R_n. The purpose of this section is to prove

THEOREM 13. *The radical of R_n is N_n.*

If M is an ideal in R, it is clear that $M \to M_n$ is a one-one mapping of the set of ideals in R onto a subset of the ideals in R_n, and that $M \subseteq M'$ if and only if $M_n \subseteq M'_n$.

LEMMA 1. *Let L be an ideal in R_n and a the element in the (i, j) position of a matrix A in the ideal L. If x and y are elements of R, then L contains a matrix with xay in the (p, q) position and zeros elsewhere.*

If $x \in R$, let X_{ij} denote the matrix with x in the (i, j) position and zeros elsewhere. Then $X_{pi}AY_{jq}$ is the required matrix in L.

LEMMA 2. *S is a simple ring with unit element if and only if S_n is a simple ring with unit element.*

If S is simple with unit element and L a nonzero ideal in S_n, let $a \neq 0$ be an element of some matrix in L. In S, the ideal $(a) = \{\Sigma x_i a y_i\} = S$, and hence by Lemma 1, if $s \in S$, there exists in L a matrix with s appearing in any given position and zeros elsewhere. Adding matrices of this type, it is clear that $L = S_n$ and it is obvious that S_n has a unit element. Conversely, the mapping $M \to M_n$ establishes that S_n is not simple if S is not, and it is well known that S_n has no unit element if S does not.

The theorem will follow readily when we have proved

LEMMA 3. *If R/M is a simple ring with unit element, then R_n/M_n is a simple ring with unit element. Conversely, if L is an ideal in R_n such that R_n/L is a simple ring with unit element, then the set M of all elements of R which appear as elements in the matrices of L is an ideal in R, $L = M_n$ and R/M is a simple ring with unit element.*

The first assertion follows from Lemma 2 and the readily proved fact that $R_n/M_n \cong (R/M)_n$.

For the converse, let M^* be the set of elements of R which appear in the $(1, 1)$ position in matrices of L. Clearly M^* is an ideal in R and $M^* \subseteq M$. We show that $M \subseteq M^*$ and hence that $M = M^*$.

Since, by hypothesis, R_n/L has a unit element, R_n contains a matrix $U = (u_{ij})$ which is the unit element of R_n modulo L. Thus

$$UXU \equiv X \quad (L) \tag{2}$$

for every matrix X in R_n. If, in particular, we choose X to be a matrix with an arbitrary element r of R in the $(1, 1)$ position and zeros elsewhere, and consider the elements in the $(1, 1)$ position of the two members of (2), we see that

$$u_{11}ru_{11} \equiv r \quad (M^*) \tag{3}$$

for every element r of R. If $b \in M$, there is a matrix in L containing b in some position. Lemma 1 then shows that there exists a matrix in L con-

taining $u_{11}bu_{11}$ in the $(1,1)$ position. Hence $u_{11}bu_{11} \in M^*$ and it follows from (3) that $b \in M^*$. Thus $M \subseteq M^*$, and M is an ideal in R.

We show next that $L = M_n$. It is evident that $L \subseteq M_n$, so it suffices to prove that $M_n \subseteq L$. If $c \in M$, then $UC_{hk}U$ can be considered as a sum of matrices of the type described in Lemma 1, hence $UC_{hk}U \in L$. From (2), it follows that $C_{hk} \in L$. But any matrix of M_n can be expressed as a sum of matrices of the form C_{hk}. Thus $M_n \subseteq L$ and therefore $L = M_n$.

Finally, since $(R/M)_n \cong R_n/M_n = R_n/L$, which is a simple ring with unit element, it follows from Lemma 2 that R/M is a simple ring with unit element. This completes the proof of Lemma 3.

It is apparent from Lemma 3 and the properties of the mapping $M \to M_n$ that the ideals M in R for which R/M is a simple ring with unit element correspond one-to-one with the ideals L in R_n for which R_n/L is a simple ring with unit element. Furthermore, it is clear that intersections of corresponding ideals also correspond. It follows from Theorem 7 that the radicals correspond, and thus the radical of R_n is N_n.

COROLLARY. *If R is a radical ring, R_n is a radical ring.*

6. Some other special cases of the general theory. In this section, we return to give some further illustrations and miscellaneous results of the general theory of 3. We recall that the mapping $a \to F(a)$ of R into the set of ideals in R is assumed to have the property that if $a \to \bar{a}$ is a homomorphism of R onto $\overline{R}$, then $F(\bar{a}) = \overline{F(a)}$.

In addition to the important special case, $a \to G(a)$, discussed in detail in the preceding two sections, we may list the following expressions for $F(a)$, all of which are ideals and meet the above stated condition. The list is by no means exhaustive but suggests some of the possibilities. In the following, α is an arbitrary integer:

$$F_\alpha(a) = \{ax - \alpha x + ya - \alpha y + \Sigma r_i a s_i - \alpha \Sigma r_i s_i\},$$
$$F_\alpha{}^l(a) = \{ax - \alpha x + \Sigma r_i a s_i - \alpha \Sigma r_i s_i\},$$
$$F_\alpha{}^r(a) = \{xa - \alpha x + \Sigma r_i a s_i - \alpha \Sigma r_i s_i\},$$
$$H_\alpha(a) = \{\Sigma r_i a s_i - \alpha \Sigma r_i s_i\},$$
$$G^*(a) = G(-a^2) = \{a^2 x + x + \Sigma r_i a^2 s_i + \Sigma r_i s_i\}.$$

In these expressions; x, y, r_i, s_i vary over the elements of R and the sums are finite.

It will be noted that the $G(a)$, used in the definition of the radical, is just $F_1{}^l(a)$.

Obviously, $F_\alpha(a) = 0$ if and only if $ax = xa = \alpha x$ for all x in R, that is, a is an α-*fier* of R.[12] Thus, as a consequence of Theorem 5, we obtain

THEOREM 14. *A subdirectly irreducible ring R has zero F_α-radical if and only if the minimal ideal J contains a nonzero α-fier of R.*

This theorem together with Theorems 1, 6 and 7 show that the F_1-radical of a ring coincides with the radical. In fact, the $F_1{}^r$-radical is also just the radical.

Consider next the mapping defined by $a \to G^*(a)$. Here $G^*(a) = 0$ if and only if $-a^2$ is a left unit of R. Theorems 4 and 5, together with the argument used in the proof of Theorem 6, establish at once the following result:

THEOREM 15. *A ring has zero G^*-radical if and only if it is isomorphic to a subdirect sum of simple rings, each with a unit element 1 and an element whose square is -1.*

The principle used in constructing $G^*(a)$ can obviously be generalized by using, in place of $-a^2$, any polynomial in a with integral coefficients and zero constant term. This process may also be applied to the other mappings as well as to G.

If R has a unit element, $F_\alpha = F_\alpha{}^l = F_\alpha{}^r = H_\alpha$, but in general they are different. We discuss, in some detail, the cases of H_α for $\alpha = 1$ and 0 respectively.

It is clear that $H_1(a) = 0$ if and only if

$$xay - xy = 0 \tag{4}$$

for all x and y in R. Thus, by Theorem 5, a subdirectly irreducible ring R has zero H_1-radical if and only if its minimal ideal J contains a nonzero element a for which (4) holds. We now consider two cases.

We suppose first that the subdirectly irreducible ring R has no right or left annihilator except zero. That is, $zR = 0$ implies $z = 0$, and $Rz = 0$ implies $z = 0$. From (4), we see that $(xa - x)R = 0$ and $R(ax - x) = 0$ for every x in R. Hence $ax = xa = x$, and a is the unit element of R. But $a \in J$, so R is a simple ring with unit element.

Suppose, on the other hand, that R has a nonzero element c such that $cR = 0$ or $Rc = 0$, and, for definiteness, assume that $cR = 0$. It follows that (c) consists of all elements of the form $nc + rc$, where n is an integer

[12] Brown and McCoy [1].

and $r \in R$. But $a \in (c)$ since $J \subseteq (c)$, hence $aR = 0$. From (4), we have that $xy = 0$ for all x, y in R, that is, that $R^2 = 0$.[13]

Conversely, in either case, a subdirectly irreducible ring has zero H_1-radical. Hence from Theorem 4 we obtain

THEOREM 16. *A ring has zero H_1-radical if and only if it is isomorphic to a subdirect sum of rings, each of which is either a simple ring with unit element or a subdirectly irreducible ring whose square is zero.*

Finally, we consider the case of H_α in which $\alpha = 0$. Here $H_0(a) = 0$ if and only if

$$xay = 0 \tag{5}$$

for all x and y in R. As in the preceding case, we only need to characterize those subdirectly irreducible rings R which have a nonzero element a of the minimal ideal J such that (5) is satisfied. Let us assume for the moment that $aR \neq 0$, hence that there exists an element z of R such that $az \neq 0$. Now the ideal (az) must contain a, hence in view of (5) we have

$$a = naz + azr,$$

where n is an integer and $r \in R$. From this, again using (5), it follows that $Ra = 0$. Hence either $Ra = 0$ or $aR = 0$. Since $(a) = J$, it follows at once that either $RJ = 0$ or $JR = 0$. It is clear that if this condition is satisfied, the subdirectly irreducible ring R has zero H_0-radical. Theorem 4 now establishes

THEOREM 17. *A ring has zero H_0-radical if and only if it is isomorphic to a subdirect sum of subdirectly irreducible rings R_i with the minimal ideal J_i such that either $R_iJ_i = 0$ or $J_iR_i = 0$.*

It is an easy consequence of Theorem 14 that in the decompositions induced by the vanishing of the F_0-radical, the rings R_i satisfy $J_iR_i = R_iJ_i = 0$. For the F_0^l-radical (F_0^r-radical), they satisfy $J_iR_i = 0$ ($R_iJ_i = 0$).

AMHERST COLLEGE.
SMITH COLLEGE.

[13] A detailed characterization of subdirectly irreducible rings R with $R^2 = 0$ would take us too far afield. However, it is not difficult to show that the additive group of such a ring is primary of rank one, as defined by Prüfer. Hence such a group is either of the type p^n or of the type p^∞ (Prüfer [1]).

BIBLIOGRAPHY.

R. Baer:

[1] "Radical ideals," *American Journal of Mathematics*, vol. 65 (1943), pp. 537-568.

G. Birkhoff:

[1] "Subdirect unions in universal algebra," *Bulletin of the American Mathematical Society*, vol. 50 (1944), pp. 764-768.

B. Brown and N. H. McCoy:

[1] "Rings with unit element which contain a given ring," *Duke Mathematical Journal*, vol. 13 (1946), pp. 9-20.

N. Jacobson:

[1] "The radical and semi-simplicity for arbitrary rings," *American Journal of Mathematics*, vol. 67 (1945), pp. 300-320.

G. Köthe:

[1] "Die Struktur der Ringe, deren Restlassenring nach dem Radikal vollständig reduzibel ist," *Mathematische Zeitschrift*, vol. 32 (1930), pp. 161-186.

J. Levitzki:

[1] "On the radical of a general ring," *Bulletin of the American Mathematical Society*, vol. 49 (1943), pp. 462-466.

[2] "Semi-nilpotent ideals," *Duke Mathematical Journal*, vol. 10 (1943), pp. 553-556.

N. H. McCoy:

[1] "Subrings of infinite direct sums," *Duke Mathematical Journal*, vol. 4 (1938), pp. 486-494.

[2] "Subdirectly irreducible commutative rings," *Duke Mathematical Journal*, vol. 12 (1945), pp. 381-387.

N. H. McCoy and D. Montgomery:

[1] "A representation of generalized Boolean rings," *Duke Mathematical Journal*, vol. 3 (1937), pp. 455-459.

S. Perlis:

[1] "A characterization of the radical of an algebra." *Bulletin of the American Mathematical Society*, vol. 48 (1942), pp. 128-132.

H. Prüfer:

[1] "Untersuchungen über die Zerlegbarkeit der abzählbaren primären Abelschen Gruppen," *Mathematische Zeitschrift*, vol. 17 (1923), pp. 35-61.

M. H. Stone:

[1] "The theory of representations for Boolean algebras," *Transactions of the American Mathematical Society*, vol. 40 (1936), pp. 37-111.

ON CONGRUENCES AND CONJUGATE NETS.*

By V. G. Grove.

---

1. Introduction. In this paper we study conjugate nets from a metric point of view. The equations of the sustaining surface, assumed to be non-developable, are written in terms of arbitrary parameters. The formulas convenient for the study of the net are written in terms of these parameters and certain normalized components of vectors along the tangents to the curves of the net. An invariant of a curve of the net, which we have called the *asymptotic curvature* of the curve, has been found which is related to isothermally conjugate nets in a manner similar to the relation between the geodesic curvature of the curves of an orthogonal net to isothermic nets. Conditions that a congruence of lines protruding from the surface be conjugate to the sustaining surface, and to the net are found. Some attention is paid to pencils of conjugate nets.

Let the parametric equations of the surface S be $x^i = x^i(u^1, u^2)$ and let $X^i(u^1, u^2)$ be the direction cosines of the normal to S at the point x with coordinates (x^1, x^2, x^3). These functions satisfy the differential equations

$$x^i{}_{,\alpha\beta} = d_{\alpha\beta}X^i, \qquad X^i{}_{,\alpha} = m_\alpha{}^\rho x^i{}_{,\rho}, \qquad m_\alpha{}^\beta = -d_{\alpha\rho}g^{\rho\beta} \tag{1.1}$$

the comma denoting covariant differentiation with respect to the metric tensor $g_{\alpha\beta}$. Let $du^1/du^2 = U^1/U^2$, $du^1/du^2 = V^1/V^2$ be the differential equations of the conjugate net N. Then $d_{\rho\sigma}U^\rho V^\sigma = 0$. We shall call the curve of N whose tangent vector is $U^\alpha(V^\alpha)$ the U-curve (V-curve) with similar names for the tangents to these curves.

Since only the ratios $U^1 : U^2$, $V^1 : V^2$ are material, these components may be normalized so that

$$d_{\rho\sigma}U^\rho U^\sigma = e_1, \qquad d_{\rho\sigma}V^\rho V^\sigma = e_2 \tag{1.2}$$

wherein e_1, e_2 are ± 1 according as the radii of normal curvature $R_{(u)}$, $R_{(v)}$ in the directions of the U- and V-tangents are positive or negative. From (1.2) and the definitions of these curvatures we have

$$R_{(u)} = e_1 g_{\rho\sigma}U^\rho U^\sigma, \qquad R_{(v)} = e_2 g_{\rho\sigma}V^\rho V^\sigma.$$

---

* Received July 9, 1946.

Since S is non-developable we may define the functions $d^{\alpha\beta}$ by the expression $d^{\alpha\rho}d_{\rho\beta} = \delta_\beta^\alpha$. Let

$$U_\alpha = d_{\alpha\rho}U^\rho, \qquad V_\alpha = d_{\alpha\rho}V^\rho.$$

The following relations exist between U^α, V^α:

$$(1.3)\quad \begin{aligned} &e_1U^\alpha U^\beta + e_2V^\alpha V^\beta = d^{\alpha\beta}, \qquad e_1U_\alpha U_\beta + e_2V_\alpha V_\beta = d_{\alpha\beta}, \\ &U^\alpha V^\beta - V^\alpha U^\beta = \epsilon^{\alpha\beta}, \qquad V^\alpha = \epsilon^{\alpha\rho}U_\rho \qquad U^\alpha = \epsilon^{\alpha\rho}V_\rho \end{aligned}$$

wherein

$$\begin{aligned} &\epsilon^{11} = \epsilon^{22} = 0, \qquad \epsilon^{12} = -\epsilon^{21} = (ed)^{-\frac{1}{2}}, \qquad e = e_1e_2 \\ &\epsilon_{11} = \epsilon_{22} = 0, \qquad \epsilon_{12} = -\epsilon_{21} = (ed)^{\frac{1}{2}}. \end{aligned}$$

2. S referred to N, and the asymptotic net. Let us now write the differential equations (1.1) of S in terms of the conjugate parameters of N. To this end put

$$(2.1)\quad x^i{}_1 = fU^\rho x^i{}_{,\rho}, \qquad x^i{}_2 = gV^\rho x^i{}_{,\rho}, \qquad F = \log f, \qquad G = \log g.$$

wherein f and g are to be chosen so that $(x^i{}_1)_2 = (x^i{}_2)_1$. We find from (2.1) that

$$x^i{}_{,\alpha} = e_1U_\alpha x^i{}_1/f + e_2V_\alpha x^i{}_2/g.$$

and, since $(x^i{}_1)_2 = (x^i{}_2)_1$

$$(2.2)\quad V^\rho F_{,\rho} = e_1U_\rho W^\rho, \qquad U^\rho G_{,\rho} = -e_2V_\rho W^\rho,$$

wherein

$$W^\alpha = U^\rho V^\alpha{}_{,\rho} - V^\rho U^\alpha{}_{,\rho}.$$

The functions x^i, X^i satisfy the equations

$$(2.3)\quad x^i{}_{\alpha\beta} = l^\rho{}_{\alpha\beta}x^i{}_\rho + D_{\alpha\beta}X^i, \qquad X^i{}_\alpha = M^\rho{}_\alpha x^i{}_\rho,$$

wherein

$$(2.4)\quad \begin{aligned} &l^1{}_{11} = e_1f(e_1U^\rho F_{,\rho} + U_\rho U^\rho{}_{,\sigma}U^\sigma), \qquad l^1{}_{22} = e_1g^2U_\rho V^\rho{}_{,\sigma}V^\sigma/f, \\ &l^2{}_{11} = e_2f^2V_\rho U^\rho{}_{,\sigma}U^\sigma/g, \qquad l^2{}_{22} = e_2g(e_2V^\rho G_{,\rho} + V_\rho V^\rho{}_{,\sigma}V^\sigma), \\ &l^1{}_{12} = e_1gU_\rho V^\rho{}_{,\sigma}U^\sigma, \qquad l^2{}_{12} = e_2fV_\rho U^\rho{}_{,\sigma}V^\sigma, \\ &D_{11} = e_1f^2, \qquad D_{12} = 0, \qquad D_{22} = e_2g^2, \\ &M^1{}_1 = -e_1h_{\rho\sigma}U^\rho U^\sigma, \qquad M^2{}_1 = -e_2fh_{\rho\sigma}U^\rho V^\sigma/g, \\ &M^2{}_1 = -e_1gh_{\rho\sigma}U^\rho V^\sigma/f, \qquad M^2{}_2 = -e_2h_{\rho\sigma}V^\rho V^\sigma, \end{aligned}$$

$h_{\alpha\beta}$ being the metric tensor of the spherical representation of S.

One may show readily that, under the conditions (2.2), the expressions

$$du = U_\rho du^\rho/f, \qquad dv = V_\rho du^\rho/g$$

are exact, and hence (2.3) are the Gauss differential equations of S in the parameters u, v and in which $x^i{}_1 = \partial x^i/\partial u$, etc.

As an illustration of the use one may make of (2.3), we recall that the focal point on the U-tangent is given by the formula

$$\rho^i{}_{(u)} = x^i - x^i{}_1/l^2{}_{12} = x^i - e_2 U^\rho x^i{}_{,\rho}/(V_\sigma U^\sigma{}_{,\lambda} V^\lambda).$$

Similarly the focal point on the V-tangent is given by

$$\rho^i{}_{(v)} = x^i - e_1 V^\rho x^i{}_{,\rho}/(U_\sigma V^\sigma{}_{,\lambda} U^\lambda). \tag{2.5}$$

Moreover the direction numbers λ^i of the axis of N are given by the formula

$$\begin{aligned}\lambda^i &= e_2 l^1{}_{22} x^i{}_1/g^2 + e_1 l^2{}_{11} x^i{}_2/f^2 + X^i \\ &= e[U_\rho V^\rho{}_{,\sigma} V^\sigma U^\lambda + V_\rho U^\rho{}_{,\sigma} U^\sigma V^\lambda] x^i{}_{,\lambda} + X^i.\end{aligned}$$

Again N is isothermally conjugate if and only if $(\log D_{11}/D_{22})_{12} = 0$. Using (2.1), (2.2) and (2.3) this condition reduces to the vanishing of the invariant I defined by

$$I = e_1 (U_\rho W^\rho)_{,\sigma} U^\sigma + e_2 (V_\rho W^\rho)_{,\sigma} V^\sigma. \tag{2.6}$$

Define the vectors A^a, B^a by the formulas

$$A^a = U^a + \sqrt{-e}\, V^a, \qquad B^a = U^a - \sqrt{-e}\, V^a. \tag{2.7}$$

One finds readily that $d_{\rho\sigma} A^\rho A^\sigma = 0$, $d_{\rho\sigma} B^\rho B^\sigma = 0$, hence these vectors are tangent vectors of the asymptotic curves. We shall speak of these tangents as the A- and B-tangents.

Defining $x^i{}_1$, $x^i{}_2$ by the expressions

$$x^i{}_1 = m A^\rho x^i{}_{,\rho}, \qquad x^i{}_2 = n B^\rho x^i{}_{,\rho}, \qquad M = \log m, \qquad N = \log n;$$

and demanding that $(x^i{}_1)_2 = (x^i{}_2)_1$ we find that m and n must satisfy the conditions

$$B^\rho M_{,\rho} = -e_1 \sqrt{-e}\, B_\rho W^\rho, \; A^\rho N_{,\rho} = e_1 \sqrt{-e}\, A_\rho W^\rho.$$

It follows that x^i, X^i are solutions of equations of the form (2.3) with coefficients defined by

$$\begin{aligned}
&l^1{}_{11} = \tfrac{1}{2} e_1 m A^\rho (2 e_1 M_{,\rho} + B_\sigma A^\sigma{}_{,\rho}), && l^1{}_{22} = \tfrac{1}{2} n^2 B_\rho B^\rho{}_{,\sigma} B^\sigma/m, \\
&l^2{}_{11} = \tfrac{1}{2} e_1 m^2 A_\rho A^\rho{}_{,\sigma} A^\sigma, && l^2{}_{22} = \tfrac{1}{2} e_1 n B^\rho (2 e_1 N_{,\rho} + A_\sigma B^\sigma{}_{,\rho}), \\
&l^1{}_{12} = \tfrac{1}{2} e_1 n B_\rho B^\rho{}_{,\sigma} A^\sigma, && l^2{}_{12} = \tfrac{1}{2} e_1 m A_\rho A^\rho{}_{,\sigma} B^\sigma, \\
&M^1{}_1 = -\tfrac{1}{2} e_1 h_{\rho\sigma} A^\rho B^\sigma, && M^2{}_1 = -\tfrac{1}{2} e_1 m h_{\rho\sigma} A^\rho A^\sigma/n, \\
&M^1{}_2 = -\tfrac{1}{2} e_1 n h_{\rho\sigma} B^\rho B^\sigma/m, && M^2{}_2 = -\tfrac{1}{2} e_1 h p_{\rho\sigma} A^\rho B^\sigma, \\
&D_{11} = D_{22} = 0, \qquad D_{12} = 2mne_1.
\end{aligned} \tag{2.8}$$

As an application, we observe that S is ruled if $R_1R_2 = 0$ and a quadric if $R_1 = R_2 = 0$ where

$$R_1 = e_1 A_\rho A^\rho{}_{,\sigma} A^\sigma, \qquad R_2 = e_2 B_\rho B^\rho{}_{,\sigma} B^\sigma. \tag{2.9}$$

Using (2.7) we may write (2.9) in the form

$$R_1 = \Phi - \sqrt{-e}\,\Psi, \qquad R_2 = \Phi + \sqrt{-e}\,\Psi,$$

wherein

$$\Phi = e_1 U_\rho U^\rho{}_{,\sigma} U^\sigma - e_2 (U_\rho V^\rho{}_{,\sigma} V^\sigma + V_\rho U^\rho{}_{,\sigma} V^\sigma + V_\rho V^\rho{}_{,\sigma} U^\sigma),$$
$$\Psi = e_2 V_\rho V^\rho{}_{,\sigma} V^\sigma - e_1 (V_\rho U^\rho{}_{,\sigma} U^\sigma + U_\rho V^\rho{}_{,\sigma} U^\sigma + U_\rho U^\rho{}_{,\sigma} V^\sigma).$$

Hence S is a quadric if and only if $\Phi = \Psi = 0$.

As a second application of this section, it is well known that the polar reciprocal l_2 if the line l_1 joining the points

$$r^i{}_{(1)} = x^i - e_1 x^i{}_1/(bm), \qquad r^i{}_2 = x^i - e_1 x^i{}_2/(an)$$

has direction numbers

$$x^i{}_{12} - e_1 anx^i{}_1 - e_1 bmx^i{}_2.$$

Using (2.8) the direction numbers of l_2 assume the form

$$\lambda^i = \tfrac{1}{4}[(B_\rho B^\rho{}_{,\sigma} A^\sigma - 2a)A^\lambda + (A_\rho A^\rho{}_{,\sigma} B^\sigma - 2b)B^\lambda]x^i{}_{,\lambda} + X^i.$$

It follows that the reciprocal of the normal to S joins the points $r^i{}_{(1)}$, $r^i{}_{(2)}$ defined by

$$r^i{}_{(1)} = x^i - 2e_1 A^\rho x^i{}_{,\rho}/(A_\sigma B^\sigma{}_{,\lambda} B^\lambda),$$
$$r^i{}_{(2)} = x^i - 2e_1 B^\rho x^i{}_{,\rho}/(B_\sigma A^\sigma{}_{,\lambda} A^\lambda).$$

3. The asymptotic curvature of a curve. Denoting covariant differentiation with respect to the tensor $d_{\alpha\beta}$ by a semicolon, we find from (1.2) that

$$d_{\rho\sigma} U^\rho U^\sigma{}_{;\lambda} U^\lambda = 0.$$

Hence the vector $U^\alpha{}_{;\rho} U^\rho$ is conjugate to the U-tangent, and we may write

$$U^\alpha{}_{;\rho} U^\rho = k_{(u)} V^\alpha. \tag{3.1}$$

Using the third of (1.3) we find that

$$k_{(u)} = \epsilon_{\rho\sigma} U^\rho U^\sigma{}_{;\lambda} U^\lambda. \tag{3.2}$$

We shall call the function $k_{(u)}$ defined by (3.1) or (3.2) *the asymptotic curvature of the U-curve.*

Similarly the asymptotic curvature of the V-curve is given by the expression

$$k_{(v)} = \epsilon_{\rho\sigma} V^\rho V^\sigma{}_{;\lambda} V^\lambda.$$

An alternate form of (3.2) may be found by multiplying both members of (3.1) by V_α and summing on α. We find that

$$k_{(u)} = e_2 V_\rho U^\rho{}_{;\sigma} U^\sigma. \tag{3.3}$$

From (1.2) we find that $U_\rho U^\rho{}_{;\sigma} V^\sigma = 0$. Hence (3.3) may be written in the form

$$k_{(u)} = e_2 U^\rho{}_{;\sigma}(V_\rho U^\sigma - U_\rho V^\sigma) = e_2 d_{\rho\lambda}\epsilon^{\sigma\lambda} U^\rho{}_{;\sigma} = e\,\epsilon^{\rho\sigma} U_{\sigma,\rho}. \tag{3.4}$$

Since

$$U_\rho W^\rho = U_\rho(U^\sigma V^\rho{}_{;\sigma} + V^\sigma U^\rho{}_{;\sigma}) = (U^\rho V^\sigma - V^\rho U^\sigma) U_{\rho,\sigma},$$

we have other expressions for $k_{(u)}$, namely

$$k_{(u)} = -e_2 U_\rho W^\rho = -eV^\rho F_{,\rho}. \tag{3.5}$$

Similarly

$$k_{(v)} = e_1 \epsilon^{\rho\sigma} V_{\sigma,\rho} = e_1 V_\rho W^\rho = -eU^\rho G_{,\rho}. \tag{3.6}$$

If the parametric curves on S are conjugate and if these curves are chosen as the U- and V-curves, then we may write

$$U^1 = (e_1 d_{11})^{-\frac{1}{2}}, \quad U^2 = 0, \quad U_1 = (e_1 d_{11})^{\frac{1}{2}}, \quad U_2 = 0,$$
$$V^1 = 0, \quad V^2 = (e_2 d_{22})^{-\frac{1}{2}}, \quad V_1 = 0, \quad V_2 = (e_2 d_{22})^{\frac{1}{2}}.$$

Then $k_{(u)}$ and $k_{(v)}$ may be written in the forms

$$k_{(u)} = -e_2 (ed)^{-\frac{1}{2}} \partial/\partial u^2 (e_1 d_{11})^{\frac{1}{2}}, \qquad k_{(v)} = -e_1 (ed)^{-\frac{1}{2}} \partial/\partial u^1 (e_2 d_{22})^{\frac{1}{2}}. \tag{3.7}$$

These latter forms of the asymptotic curvatures of the curves of a conjugate net show their resemblance to the geodesic curvature of orthogonal curves.

Comparing (3.5) and (3.6) with (2.6) we find that *if the asymptotic curvatures of the curves of a conjugate net are constant along the corresponding curves of the net, the net is isothermally conjugate.*

The curve whose differential equation is $du^1/du^2 = C^1/C^2$ is an extremal [1] of the integral

$$I = \int \sqrt{e d_{\rho\sigma} du^\rho du^\sigma}$$

if the components C^α satisfy the equation

$$\epsilon_{\rho\sigma} C^\rho C^\sigma{}_{;\lambda} C^\lambda = 0.$$

It follows from (3.2) that *a non-asymptotic curve is an extremal of the integral I only if the asymptotic curvature of the curve vanishes at each of its points.*

Now define the vectors $\bar{U}^a$, $\bar{V}^a$ by the formulas

$$\begin{aligned} \bar{U}^a &= U^a \cos\theta + V^a \sin\theta, \\ \bar{V}^a &= -U^a \sin\theta + V^a \cos\theta, \end{aligned} \tag{3.8}$$

if $e_1 = e_2$, $e = 1$, that is, if $K > 0$; and by the formulas

$$\begin{aligned} \bar{U}^a &= U^a \cosh\theta + V^a \sinh\theta, \\ \bar{V}^a &= U^a \sinh\theta + V^a \cosh\theta, \end{aligned} \tag{3.9}$$

if $e_1 = -e_2$, $e = 1$, that is, if $K < 0$. In either case $d_{\rho\sigma}\bar{U}^\rho\bar{U}^\sigma = e_1$, $d_{\rho\sigma}\bar{V}^\rho\bar{V}^\sigma = e_2$, and hence these vectors are normalized. One may readily show that the $\bar{U}$- and $\bar{V}$-tangents are conjugate, and that the curves defined by them form a pencil of conjugate nets if and only if $\theta = \text{const.}$ Denoting the asymptotic curvatures of the $\bar{U}$-curves and $\bar{V}$-curves by $\bar{k}_{(u)}$ and $\bar{k}_{(v)}$ we find from (3.4) and (3.6) that

$$\begin{aligned} \bar{k}_{(u)} &= k_{(u)} \cos\theta - k_{(v)} \sin\theta, \\ \bar{k}_{(v)} &= k_{(u)} \sin\theta + k_{(v)} \cos\theta, \quad K > 0; \end{aligned}$$

and

$$\begin{aligned} \bar{k}_{(u)} &= k_{(u)} \cosh\theta + k_{(v)} \sinh\theta, \\ \bar{k}_{(v)} &= k_{(u)} \sinh\theta + k_{(v)} \cosh\theta, \quad K < 0. \end{aligned}$$

For these two cases therefore we observe that

$$\begin{aligned} \bar{k}^2_{(u)} + \bar{k}^2_{(v)} &= k^2_{(u)} + k^2_{(v)}, \quad K > 0, \\ \bar{k}^2_{(u)} - \bar{k}^2_{(v)} &= k^2_{(u)} - k^2_{(v)}, \quad K < 0. \end{aligned}$$

In either case we observe that *if the curves of a conjugate net are the extremals of the integral I, then all curves of the pencil determined by the net are also extremals of I; moreover the given net is isothermally conjugate.*

If we denote the transform of the invariant I defined by (2.6) under the transformations (3.8) with $\theta = \text{const.}$ by $\bar{I}$, we find that $\bar{I} = I$. We therefore corroborate Wilczynski's theorem [3] that *if one net of a pencil of conjugate nets is isothermally conjugate, all nets of the pencil are isothermally conjugate.*

If $K > 0$, the associate conjugate net of the given net N is given by (3.8) with $\theta = \pi/4$, and if $K < 0$ by (3.9) with $\theta = \pi i/4$. Suppose N is such that its associate net form the lines of curvature. Then from

$g_{\rho\sigma}\bar{U}^\rho\bar{V}^\sigma = 0$, $\theta = \pi/4$, we find that the radii of normal curvature of the curves of N are given by

$$R_{(u)} = R_{(v)} = \tfrac{1}{2}(R_1 + R_2)$$

R_1, R_2 being the principal radii of normal curvature. *The net whose associate net is the lines of curvature is then the net of mean curvature.* The same result follows from (3.9), it being observed that the given net is then imaginary, the lines of curvature being real. The radii of normal curvature in the directions of the tangents to the curves of any net (3.8) whose basis net is the net of mean curvature are given by

$$\bar{R}_{(u)} = R_1 \cos^2\left(\theta - \frac{\pi}{4}\right) + R_2 \sin^2\left(\theta - \frac{\pi}{4}\right),$$

$$\bar{R}_{(v)} = R_1 \sin^2\left(\theta - \frac{\pi}{4}\right) + R_2 \cos^2\left(\theta - \frac{\pi}{4}\right).$$

From (2.4) we observe that the U-curve is a geodesic if $l^2_{11} = 0$, that is if

$$V_\rho U^\rho{}_{,\sigma} U^\sigma = 0.$$

Moreover its asymptotic curvature is zero if and only if

$$V_\rho U^\rho{}_{;\sigma} U^\sigma = 0. \tag{3.10}$$

If $L^\alpha_{\beta\gamma}$ and $\Lambda^\alpha_{\beta\gamma}$ are respectively the Christoffel symbols formed from the tensors $g_{\alpha\beta}$, $d_{\alpha\beta}$, one finds readily that

$$\Lambda^\alpha_{\beta\gamma} = L^\alpha_{\beta\gamma} + T^\alpha_{\beta\gamma}$$

wherein

$$T^\alpha_{\beta\gamma} = \tfrac{1}{2} d^{\alpha\rho} d_{\beta\gamma,\rho}. \tag{3.11}$$

If the U-curve is a geodesic, equation (3.10) may be written in the form

$$T^\rho{}_{\sigma\lambda} V_\rho U^\sigma U^\lambda = 0$$

which by (3.11) may be written in the form

$$U_\rho V^\rho{}_{,\sigma} U^\sigma = 0. \tag{3.12}$$

Comparing (3.12) with (2.5) we observe that *if the U-curve is a geodesic, it has zero asymptotic curvature if and only if the ray of the net is parallel to the V-tangent.*

The meridians on a surface of revolution are known to be geodesics. We shall find the condition that their asymptotic curvature vanish. If the equations of S are written in the form

$$x^1 = u^1 \cos u^2, \qquad x^2 = u^1 \sin u^2, \qquad x^3 = \phi(u^1),$$

then

$$d_{11} = \phi''/(1 + (\phi')^2)^{\frac{1}{2}}, \quad d_{12} = 0, \quad d_{22} = u^1\phi'/(1 + (\phi')^2)^{\frac{1}{2}}, \quad \phi' = d\phi/du^1.$$

Hence from (3.7) the asymptotic curvature of the parallels vanishes; the asymptotic curvature of the meridians vanishes if and only if

$$\partial/\partial u^1 \log \{(u^1\phi')^2/1 + (\phi')^2\} = 0.$$

Integrating we find that

$$\phi = c \cosh^{-1} u^1/c.$$

Hence *the only surfaces of revolution whose meridians have zero asymptotic curvatures are the catenoids*; the parallels and meridians of this surface therefore form an isothermally conjugate net, a well known fact.

4. Congruences conjugate to S and N. Let there be given a congruence λ of lines l protruding from S at x. The components of the vector along l may be written in the form

$$\lambda^i = p(\theta^\rho x^i{}_{,\rho} + X^i), \qquad p \neq 0, \qquad P = \log p. \tag{4.1}$$

From (4.1) we find

$$\lambda^i{}_{,\sigma} = \mu^\rho{}_\sigma x^i{}_{,\rho} + (P_{,\sigma} + \theta_\sigma)\lambda^i, \tag{4.2}$$

wherein

$$\mu^\rho{}_\sigma = p(\theta^\rho{}_{,\sigma} + m^\rho{}_\sigma - \theta_\sigma\theta^\rho).$$

The curves on S corresponding to the developables of λ are the integral curves of the differential equation

$$\epsilon_{\rho\lambda}\mu^\rho{}_\sigma du^\lambda du^\sigma = 0.$$

It follows that λ is conjugate to S if and only if

$$\epsilon_{\rho\lambda}\mu^\rho{}_\sigma d^{\lambda\sigma} = 0. \tag{4.3}$$

Using the fact that $\epsilon^{\alpha\beta}$ is skew symmetric, and that $d_{\alpha\beta,\gamma} - d_{\alpha\gamma,\beta} = 0$, condition (4.3) may be reduced [1] to

$$\theta_{\alpha,\beta} - \theta_{\beta,\alpha} = 0. \tag{4.4}$$

Condition (4.4) is equivalent to the condition that the differential equation $\theta_\rho du^\rho = 0$ is an exact differential equation. Let therefore $\omega(u^1, u^2) = \text{const.}$ be an arbitrary one-parameter family of curves whose

differential equation is $\omega_{,\rho}du^\rho = 0$. A point y on the conjugate to the curve C of the family through x has coordinates defined by the expression

$$y^i = x^i + pd^{\rho\sigma}\omega_{,\sigma}x^i{}_{,\rho}.$$

Erect at y a perpendicular of length cp ($c = \text{const.}$) to the tangent plane to S, determining a point z. The direction numbers of the line l joining x to z are of the form

$$\lambda^i = p(d^{\rho\sigma}\omega_{,\sigma}x^i{}_{,\rho} + cX^i).$$

The most general congruence λ conjugate to S is generated by this line l.

A special case gives rise to an interesting class of congruences conjugate to S. Suppose in (4.1) that $d_{\rho\sigma}\theta^\rho\theta^\sigma = e_1$. The condition (4.4) that λ be conjugate to S, is equivalent to saying that the θ-curve has zero asymptotic curvature. Moreover this condition is invariant under the transformation $\theta_a = c\bar{\theta}_a$ ($c = \text{const}$). This class of congruences may be constructed as follows: *Select an arbitrary one parameter family of extremals of the integral I; at the intersections of the tangent lines to these curves with the Dupin indicatrix, erect perpendiculars of constant length to the tangent planes to S. The lines l determined by the points x of S and the terminal points of these perpendiculars generate a congruence conjugate to S.*

From (4.2) one finds that the conditions[2] that the developables of λ correspond to a parametric net are

$$\mu^\alpha{}_\beta = 0, \qquad \alpha \neq \beta.$$

Multiply (4.2) by fU^σ and then by gV^σ and sum on σ. In the notation of (2.1), the resulting equations may be written in the form

$$\lambda^i{}_a = \nu^\rho{}_a x^i{}_\rho + A_a\lambda^i$$

wherein

$$\nu_1{}^2 = e_2 f\mu^\rho{}_\sigma U^\sigma V_\rho/g, \qquad \nu_2{}^1 = e_1 g\mu^\rho{}_\sigma V^\sigma U_\rho/f,$$

the remaining coefficients not being material for our purposes. *The developables of λ therefore correspond to the U- and V-curves if and only if*

$$\mu^\rho{}_\sigma U^\sigma V_\rho = 0, \qquad \mu^\rho{}_\sigma V^\sigma U_\rho = 0. \tag{4.5}$$

Conditions (4.5) are not independent; in fact the vanishing of their sum is implied by (4.4). If in the first of (4.5) we let $\theta^\rho = -d^{\rho\lambda}(\log\omega)_{,\lambda}$ that equation may be written in the equivalent forms

$$U^\sigma V^\lambda(\omega_{,\sigma\lambda} - 2T^\rho{}_{\sigma\lambda}\omega_{,\rho} - h_{\sigma\lambda}\omega) = 0,$$
$$U^\sigma V_\rho[(d^{\rho\lambda}\omega_{,\lambda})_{,\sigma} + d_{\sigma\lambda}g^{\rho\lambda}\omega] = 0.$$

This latter form is equivalent to Springer's result.[2]

5. Normal congruences. There is a striking similarity between the theory of congruences conjugate to a surface S, and that of congruences normal to a second surface S'. In (4.1) choose p so that $\lambda^i\lambda^i = 1$, and let $\phi^\alpha = p\theta^\alpha$. Then (4.1) assumes the form

$$\lambda^i = \phi^\rho x^i{}_{,\rho} + pX^i, \qquad \lambda^i\lambda^i = 1 = p^2 + g_{\rho\sigma}\phi^\rho\phi^\sigma.$$

We may write the parametric equations of any transversal surface S' of λ in the form

$$y^i = x^i + q\lambda^i.$$

Then λ is the normal congruence of S' if $\lambda^i y^i{}_{,\alpha} = 0$ for $\alpha = 1, 2$. These conditions are equivalent to the conditions

$$g_{\alpha\rho}\phi^\rho + q_{,\alpha} = 0.$$

Hence λ is normal to S' if and only if

$$(g_{\alpha\rho}\phi^\rho)_{,\beta} - (g_{\beta\rho}\phi^\rho)_{,\alpha} = 0. \tag{5.1}$$

Let $\phi_\alpha = g_{\alpha\rho}\phi^\rho$; then (5.1) becomes [2]

$$\phi_{\alpha,\beta} - \phi_{\beta,\alpha}. \tag{5.2}$$

Let $\omega(u^1, u^2) = c$ be a one-parameter family of curves whose differential equation is $\omega_{,\rho}du^\rho = 0$. A point y on the tangent to S perpendicular to the tangent to the curve C of this family through x has coordinates

$$y^i = x^i + cg^{\rho\sigma}\omega_{,\sigma}x^i{}_{,\rho},$$

wherein c is a constant such that $c^2g^{\rho\sigma}\omega_{,\rho}\omega_\sigma < 1$. At y erect a perpendicular to the tangent plane to S at x intersecting the unit sphere with center at x in the points z_1, z_2. *The line l determined by x and z_1 (or z_2) generates the most general normal congruence.*

By analogy to the method in the latter part of **4**, consider a vector ϕ^α of constant length c, that is let

$$g_{\rho\sigma}\phi^\rho\phi^\sigma = c^2.$$

We find that

$$g_{\rho\sigma}\phi^\rho\phi^\sigma{}_{,\lambda}\phi^\lambda = 0.$$

The vector $\phi^\alpha{}_{,\rho}\phi^\rho$ is therefore perpendicular to ϕ^α. Then

$$\phi^\alpha{}_{,\rho}\phi^\rho = c^2 k_g \psi^\alpha, \qquad g_{\rho\sigma}\psi^\rho\psi^\sigma = 1, \qquad g_{\rho\sigma}\phi\psi^\sigma = 0$$

where k_g is the geodesic curvature of the ϕ-curve, k_g being given by

$$k_g = \sqrt{eK}\,\epsilon^{\rho\sigma}\phi_{\rho,\sigma}, \qquad \phi_\alpha = g_{\alpha\rho}\phi^\rho.$$

Hence the condition (5.2), for a vector of constant length is equivalent to the vanishing of k_i Our results may be stated as follows: *Select a one parameter family of geodesics on S, lay off points on these tangents at a constant distance from the points of contact; erect perpendiculars to the tangent planes of S of constant length from these points. The lines l determined by x and the terminal points of these segments generate a congruence normal to some surface S′.*

MICHIGAN STATE COLLEGE.

REFERENCES.

1. V. G. Grove, "A general theory of surfaces and conjugate nets," *Transactions of the American Mathematical Society*, vol. 57 (1945), pp. 105-122.
2. C. E. Springer, "Rectilinear congruences whose developables intersect a surface in its lines of curvature," *Bulletin of the American Mathematical Society*, vol. 51 (1945), pp. 990-996.
3. E. J. Wilczynski, "Geometrical significance of isothermal conjugacy," *American Journal of Mathematics*, vol. 42 (1920), pp. 211-221.

THE ELLIPTIC MODULAR FUNCTION AND A CLASS OF ANALYTIC FUNCTIONS FIRST CONSIDERED BY HURWITZ.*

By ZEEV NEHARI.

Introduction. In spite of its wide applicability in various branches of the theory of functions, the elliptic modular function is often used with a certain hesitation. This is mainly due to the fact that its application presupposes familiarity with a comparatively intricate formalism, in particular when the determination of numerical constants is involved. In fact, the endeavor to avoid the elliptic modular function has given rise to an extensive mathematical literature aiming at proving certain theorems in an "elementary" way, the word "elementary" being used here as a synonym for "without making use of the elliptic modular function." As an impressive example, Picard's theorem on integral functions might be quoted.

The difficulties which beset the numerical treatment of the elliptic modular function go essentially back to the fact that, on the one hand, the formalism of this function can only be developed with the help of the Jacobian elliptic functions while, on the other hand, what is needed in the applications are the conformal mapping properties of the modular function, and the connection between these two different aspects of the modular function has to be established through the medium of the theory of Schwarz' differential parameter or by a very detailed study of the periodic properties of the Jacobian elliptic functions.

The object of the first part of this paper is to show how those properties of the elliptic modular function which are required for the applications may be derived in a simple way by the exclusive use of elementary principles of the theory of conformal representation. It will be shown that once the "modular surface" is defined, the functional equation

$$J(z) = 4\sqrt{J(z^2)}/[1+\sqrt{J(z^2)}]^2 \tag{1}$$

satisfied by the elliptic modular function can be obtained almost by inspection. This functional equation—which is a special case of Landen's transformation of the Jacobian elliptic functions—has no other solution except the elliptic

* Received May 3, 1946.

modular function, if the further conditions $J(0) = 0$, $J'(0) > 0$ and $J(z)$ regular for $|z| < 1$, are stipulated, as then the Taylor coefficients of $J(z)$ can be successively computed with the help of (1). It will then be shown that the infinite product representing the square of the modulus of the Jacobian elliptic functions (taken as a function of the period ratio) likewise satisfies equation (1) and the additional conditions. Equation (1) having only one such solution, this implies the identity of $J(z)$ and the square of the Jacobian modulus and yields, at the same time, the convenient product expansion of the modular function.

It will further be shown that, apart from (1), $J(z)$ also satisfies some other functional relations which together amount to the fact that $J(e^{\pi i z})$ is an automorphic function of x, i. e., it is invariant with respect to a certain group of linear transformations of x.

In the second part of this paper, a number of theorems on functions $f(z) = a_1 z + a_2 z^2 + \cdots$, regular for $|z| < 1$ and not vanishing there except at $z = 0$, are proved with the help of the elliptic modular function.

PART I.

1. The elliptic modular function $w = M(z)$ (the relationship between $M(z)$ and the function $J(z)$ referred to in the Introduction will appear later) may be defined as the analytic function effecting the conformal representation of a curvilinear triangle with zero angles in the z-plane on the half-plane $J_m\{w\} > 0$, the sides of the triangle being circular arcs orthogonal to the circle $|z| = 1$; the vertices of the triangle (which necessarily are situated on $|z| = 1$) may be made to correspond to the points $w = 0, 1, \infty$. By the well-known procedure of inverting the triangle along its sides, then again inverting the resulting figure along its sides etc. etc., the whole circle $|z| < 1$ will eventually be covered by an infinity of triangles. By Schwarz' symmetry principle, all these triangles are mapped by $w = M(z)$ on half-planes $J_m\{w\} > 0$ or $J_m\{w\} < 0$ which are connected with their neighboring half-planes along the stretch $0 < w < 1$ and the rays $-\infty < w < 0$ and $1 < w < \infty$, respectively. The full circle $|z| < 1$ is mapped by $w = M(z)$ on a Riemann surface, say R_1, which covers the w-plane an infinity of times, with the exception of the points $w = 0, 1, \infty$ which are isolated boundary points of R_1.

In many applications it is more convenient to use a Riemann surface slightly different from R_1. This surface, say R_2, differs from R_1 by the fact that while $w = 0, 1, \infty$ are all outside R_1, one sheet of R_2 covers the point

$w = 0$. The function mapping $|z| < 1$ on R_2 will be denoted by $w = J(z)$. In virtue of Riemann's fundamental theorem, $J(z)$ is uniquely determined by the further conditions $J(0) = 0$ and $J'(0) > 0$.

Both $M(z)$ and $J(z)$ are called 'elliptic modular functions'; indeed they are connected by a simple relation. It is seen without difficulty that the function

$$v = u(z) = \beta \cdot e^{-\alpha[(1+\gamma z)/(1-\gamma z)]} \qquad [\alpha > 0, |\beta| = |\gamma| = 1]$$

maps $|z| < 1$ on a surface which covers the circle $|v| < 1$ an infinity of times and has inside $|v| = 1$ the only boundary point $v = 0$, which moreover is a boundary point for all sheets of the surface. Therefore, if we form the function $J[u(z)]$, we see that the point $w = 0$ is "removed" from R_2, R_2 thereby being transformed into R_1. Hence

$$M(z) = J[\beta \cdot e^{-\alpha[(1+\gamma z)/(1-\gamma z)]}].$$

2. We shall now derive the fundamental functional equation of $J(z)$. To this end we consider the function $w = \sqrt{J(z^2)}$ which, as is easily seen, maps $|z| < 1$ on a surface whose only boundary points are $w = 0, \pm 1, \infty$, all of which are outside the surface, apart from $w = 0$ which is covered by one sheet only (for $z = 0$). As in the case of R_2, there are no inner branch-points to this surface. We now consider the function

$$w^* = p(z) = 4\sqrt{J(z^2)}/[1 + \sqrt{J(z^2)}]^2.$$

The critical points $w = -1, 0, 1, \infty$ are transformed by $w^* = 4w/(1+w)^2$ into $w^* = \infty, 0, 1, 0$ respectively. Hence, $w^* = p(z)$ takes the value $w^* = 0$ only for $z = 0$ and leaves out the values $w^* = 1, \infty$. Since $dw^*/dw = 4(1-w)/(1+w)^3$ vanishes only for $w = 1$ and this value is not taken by $w = \sqrt{J(z^2)}$, we obtain the result that $w^* = p(z)$ maps $|z| < 1$ on a surface R_3 whose only boundary points are $w^* = 0, 1, \infty$ which are all outside R_3, apart from $w^* = 0$ which is covered by one sheet of R_3 (for $z = 0$); moreover, R_3 contains no inner branch-points.

But this means that R_3 is identical with R_2, as there is only one Riemann surface of this description. For let $w = J(z) = a_1 z + a_2 z^2 + \cdots$ be the function mapping $|z| < 1$ on R_2 and $w = p(z) = b_1 z + b_2 z^2 + \cdots$ ($a_1 > 0$, $b_1 > 0$) the function mapping $|z| < 1$ on R_3. Clearly both functions $w(z) = J^{-1}[p(z)]$ and $w_1(z) = p^{-1}[J(z)]$ are regular for $|z| < 1$ and we have there $|w(z)| < 1$ and $|w_1(z)| < 1$, as no path within R_2 can lead out of R_3 and vice versa, both surfaces having only the isolated boundary points $w = 0, 1, \infty$ and no branch-points. The expansions of $w(z)$ and $w_1(z)$

beginning with $w(z) = (b_1/a_1)z + \cdots$ and $w_1(z) = (a_1/b_1)z + \cdots$ respectively, we have $|\, b_1/a_1 \,| \leq 1$ and $|\, a_1/b_1 \,| \leq 1$, which, a_1 and b_1 being positive, implies $a_1 = b_1$ and consequently $w(z) \equiv z$ and $w_1(z) \equiv z$. Therefore, if we choose the positive branch of $\sqrt{J(z^2)}$, we are lead to the equation $J(z) = p(z)$, i. e.,

$$J(z) = 4\sqrt{J(z^2)}/[1 + \sqrt{J(z^2)}]^2. \tag{1}$$

As can be easily verified, this functional equation permits the successive computation of the coefficients a_n of the Taylor expansion $J(z) = \sum_{n=1}^{\infty} a_n z^n$. For the first coefficient a_1 we obtain $a_1 = 4\sqrt{a_1}$, i. e., $a_1 = 16$. This coefficient determines the constant in a classical theorem by Hurwitz to be mentioned later.

3. In order to find a convenient analytical expression for $J(z)$ and, at the same time, to establish the identity of $J(z)$ with the square of the modulus of the Jacobian elliptic functions, we proceed as follows:

Let $w = f(z;\rho)$ be the analytic function, real for real values of z, which maps the annulus $\rho < |\, z \,| < 1$ on the full circle $|\, w \,| < 1$ cut along the stretch $0 \leqq w \leqq \sqrt{\zeta(\rho)}$. It can then be easily shown that the formalism connecting $\zeta(\rho)$ with $\zeta(\rho^2)$ is the same as that connecting $J(z)$ with $J(z^2)$. In fact, by inverting the annulus $\rho < |\, z \,| < 1$ along its outer boundary we find, by Schwarz' symmetry principle, that $w = f(z;\rho)$ maps the annulus $\rho < |\, z \,| < 1/\rho$ on the full w-place cut along the stretch $0 \leq w \leq \sqrt{\zeta(\rho)}$ and the ray $1/\sqrt{\zeta(\rho)} \leqq w \leqq \infty$. Remembering that the function

$$w = 4/\sqrt{\zeta(\rho)} \cdot x/(1+x)^2$$

maps the circle $|\, x \,| < 1$ on the full w-plane cut along the ray $1/\sqrt{\zeta(\rho)} \leq w \leq \infty$, we have consequently

$$4f(\rho z;\rho^2)/\sqrt{\zeta(\rho)}\,[1 + f(\rho z;\rho^2)]^2 = f(z;\rho).$$

For $z = \rho$, this reduces to

$$4f(\rho^2;\rho^2)/[1 + f(\rho^2;\rho^2]^2 = \sqrt{\zeta(\rho)} \cdot f(\rho;\rho).$$

But $f(z;\rho)$ being real for real values of z, we have necessarily

$$f(\rho;\rho) = \sqrt{\zeta(\rho)}$$

and we thus obtain for $\zeta(\rho)$ the functional equation

$$4\sqrt{\zeta(\rho^2)}/[1+\sqrt{\zeta(\rho^2)}]^2=\zeta(\rho),$$

i. e., equation (1).

Obviously, equation (1) is not only satisfied by the modular function $J(\rho)$ but also by all functions $J(\rho^\gamma)$, where γ is any complex number. In order to find out to which of these functions $\zeta(\rho)$ corresponds, we need some more information about the analytical behavior of $\zeta(\rho)$ (which so far is only defined for positive values of ρ) in the unit circle. To this end we shall construct an explicit expression for $f(z;\rho)$ which, for $z=\rho$, will furnish an expression for $\zeta(\rho)$.

Since $w=f(z;\rho)=f(z)$ maps the circle $|z|=1$ on the circle $|w|=1$ and the circle $|z|=\rho$ on part of the real axis we have, by Schwarz' symmetry principle,

$$f(1/\bar{z})=1/\overline{f(z)}, \qquad f(\rho^2/\bar{z})=\overline{f(z)}.$$

Since $f(z)$ is real for real z, we have the additional relation $f(\bar{z})=f(z)$ which, combined with the two previous ones, yields

(2a) $$f(1/z)=1/f(z)$$

and

(2b) $$f(\rho^2/z)=f(z).$$

(2a) and (2b) entail the further relation

(2c) $$f(\rho^4 z)=f(z);$$

indeed we have

$$f(\rho^4 z)=f(1/\rho^2 z)=1/f(\rho^2 z)=1/f(1/z)=f(z).$$

Accordingly, $g(x)=f(e^x)$ is an elliptic function possessing the two elementary periods $2\pi i$ and $4\log\rho$.

$f(z)$ being real for real values of z, we have $f(-\rho)=0$; moreover, $w=0$ being the end of the cut $0\leq w\leq\sqrt{\zeta(\rho)}$, it follows that $z=-\rho$ is a double zero of $f(z)$. By (2b) and (2c), all further zeros of $f(z)$ are also double and coincide with the points $-\rho^{4n+1}$ $(n=0,\pm 1,\pm 2,\cdots)$; obviously, there are no other zeros of $f(z)$ apart from these. By (2a), all poles of $f(z)$ are also double and coincide with the numbers $-\rho^{4n-1}$.

We now consider the infinite product

(3) $$\phi(z)=z\cdot\frac{\prod_0^\infty(1+\rho^{4n+1}z^{-1})^2\prod_1^\infty(1+\rho^{4n-1}z)^2}{\prod_0^\infty(1+\rho^{4n-1}z^{-1})^2\prod_1^\infty(1+\rho^{4n+1}z)^2}, \qquad (0<\rho<1),$$

which converges for all values of z differing from 0 and ∞. By its construction, $\phi(z)$ has the same zeros and poles as $f(z)$. It is further easily verified that $\phi(z)$ satisfies the equation $\phi(\rho^4 z) = \phi(z)$, i. e., $\phi(e^x)$ is an elliptic function of x possessing the same zeros, poles and elementary periods as $f(e^x)$. By Liouville's theorem, $\phi(z)$ and $f(z)$ are therefore identical, apart from a multiplicative constant. In view of (2a), this constant can only take the values ± 1. $f(z)$ being assumed positive for positive values of z, the negative sign is ruled out and we have

$$f(z) = \phi(z).$$

Remembering that $f(\rho) = \sqrt{\zeta(\rho)}$, we obtain from (3):

$$\sqrt{\zeta(\rho)} = f(\rho) = \phi(\rho) = 4\rho \prod_{n=1}^{\infty} [(1+\rho^{4n})/(1+\rho^{4n-2})]^4,$$

$$\zeta(\rho) = 16\rho^2 \prod_{n=1}^{\infty} [(1+\rho^{4n})/(1+\rho^{4n-2})]^8.$$

This product converges absolutely for all values $|\rho| < 1$ and represents there an analytic function of ρ. As shown above, this function $\zeta(\rho)$ satisfies for positive values of ρ the equation

$$\zeta(\rho) = 4\sqrt{\zeta(\rho^2)}/[1+\sqrt{\zeta(\rho^2)}]^2.$$

As a relation between analytic functions this equation must remain true for all values ρ for which the functions involved may be continued analytically, i. e., for all values $|\rho| < 1$.

As mentioned further above, the functional equation for $\zeta(\rho)$ is also satisfied by all functions $\zeta(\rho^\gamma)$. If we denote $\zeta(\sqrt{\rho})$ by $\eta(\rho)$, we have for $\eta(\rho)$ the expression

$$\eta(\rho) = 16\rho \prod_{n=1}^{\infty} [(1+\rho^{2n})/(1+\rho^{2n-1})]^8,$$

which shows that $\eta(\rho)$ is analytic for $|\rho| < 1$ and satisfies $\eta(0) = 0$, $\eta'(0) > 0$. $\eta(\rho)$ further satisfies equation (1). As there is only one such solution to equation (1), we must necessarily have

$$\eta(\rho) = J(\rho),$$

whence

$$J(z) = 16z \prod_{n=1}^{\infty} [(1+z^{2n})/(1+z^{2n-1})]^8, \tag{4}$$

which is the well known formula for the square of the modulus of the Jacobian elliptic functions with $z = e^{\pi i \tau}$, where τ is the period ratio of these functions.

4. In addition to the functional equation (1), there are three more functional equations satisfied by $J(z)$, viz.,

$$J(-z) = J(z)/(J(z)-1), \qquad |z| < 1, \tag{5}$$

$$J(e^{-\pi x}) + J(e^{-\pi/x}) = 1, \qquad R\{x\} > 0, \tag{6}$$

$$J(-e^{-\pi x}) \cdot J(-e^{-\pi/x}) = 1. \tag{7}$$

These equations are simple consequences of the fact that the set of points $(0, 1, \infty)$ is transformed into itself by the transformations $w^* = w/(w-1)$, $w^* = 1-w$ and $w^* = 1/w$ respectively.

As shown above, $w = J(z)$ maps $|z| < 1$ on a locally "schlicht" surface R_2 the only boundary points of which are $w = 0, 1, \infty$, all of them outside R_2 apart from $w = 0$ which is covered by one sheet of R_2 (for $z = 0$). This surface is obviously transformed into itself by the transformation $w^* = w/(w-1)$ which leaves the origin fixed and interchanges the two points 1 and ∞. The function mapping $|z| < 1$ on this surface is

$$w^* = J(z)/(J(z)-1) = -J'(0)z + \cdots.$$

As seen earlier, a function $f(z)$ mapping $|z| < 1$ on R_2 is completely determined by the conditions $f(0) = 0$ and $f'(0) > 0$. As both $-J(z)/(J(z)-1)$ and $-J(-z)$ satisfy these conditions, we must, therefore, have

$$J(-z) = J(z)/(J(z)-1),$$

i.e., equation (5). This relation may also be derived directly from (1). From

$$J(z) = 4\sqrt{J(z^2)}/[1+\sqrt{J(z^2)}]^2$$

follows

$$J(-z) = -4\sqrt{J(z^2)}/[1-\sqrt{J(z^2)}]^2,$$

$\sqrt{J(z^2)}$ being an odd function of z. By combining these two formulae we obtain

$$\begin{aligned} J(z)/J(-z) &= -[(1-\sqrt{J(z^2)})/(1+\sqrt{J(z^2)})]^2 \\ &= 4\sqrt{J(z^2)}/[1+\sqrt{J(z^2)}]^2 - 1 = J(z) - 1, \end{aligned}$$

i.e., (5).

In order to prove (6), we pass from R_2 to the surface R_1 described earlier, by the transformation

$$w = M(z) = J[e^{-\alpha([1+z]/[1-z])}], \qquad \alpha > 0.$$

The surface R_1 on which $|z| < 1$ is mapped by $w = M(z)$ is locally "schlicht" and has no boundary points except $w = 0, 1, \infty$, all of which are outside R_1. We now consider the function

$$w^* = \psi(z) = M(z) - \tfrac{1}{2} = J[e^{-\alpha([1+z]/[1-z])}] - \tfrac{1}{2},$$

where α ($\alpha > 0$) has been so chosen as to make $J(e^{-\alpha}) = \frac{1}{2}$. With this value of α we have $\psi(0) = 0$; besides, the surface on which $|z| < 1$ is mapped by $w^* = \psi(z)$ is symmetrical with regard to $w^* = 0$ (its boundaries being the points $w^* = \pm \frac{1}{2}, \infty$). We may, therefore, conclude that $\psi(z)$ is an odd function of z, i. e., $\psi(-z) = -\psi(z)$, or

$$(8) \qquad J[e^{-\alpha([1+z]/[1-z])}] - \tfrac{1}{2} = -J[e^{-\alpha([1-z]/[1+z])}] + \tfrac{1}{2}.$$

With $x = (1+z)/(1-z)$ ($R\{x\} > 0$) this becomes

$$(8a) \qquad J(e^{-\alpha x}) + J(e^{-\alpha/x}) = 1,$$

which is identical with (6) provided we have $J(e^{-\pi}) = \frac{1}{2}$ (as will be shown presently).

Equation (7) is obtained by combining (5) and (8a). Indeed, by (5):

$$J(-e^{-\alpha x}) \cdot J(-e^{-\alpha/x}) = J(e^{-\alpha x})J(e^{-\alpha/x})/[J(e^{-\alpha x}) - 1][J(e^{-\alpha/x}) - 1].$$

In view of (8a), the terms on the right-hand side cancel out and we obtain

$$(8b) \qquad J(-e^{-\alpha x}) \cdot J(-e^{-\alpha/x}) = 1.$$

It now remains to show that $J(e^{-\pi}) = \frac{1}{2}$. To this end we note that, by (5), $J(e^{-\alpha}) = \frac{1}{2}$ entails $J(-e^{-\alpha}) = -1$. By virtue of (1) we have further

$$J(-ie^{-\alpha/2}) = 4\sqrt{J(-e^{-\alpha})}/[1 + \sqrt{J(-e^{-\alpha})}]^2 = 4i/(1+i)^2 = 2,$$

whence

$$J(e^{-\alpha})J(-ie^{-\alpha/2}) = 1.$$

Although formula (8b) does not generally mean that for any two values z_1 and z_2 for which $J(z_1) \cdot J(z_2) = 1$ there exists a value x such that $z_1 = -e^{-\alpha x}$ and $z_2 = -e^{-\alpha/x}$, we may show that in our case this is nevertheless true. In view of $J(-e^{-\alpha}) = -1$, (8a) is fulfilled for $x = 1$. We now connect $w = -1$ with $w = \frac{1}{2}$ by a path t in the upper half-plane and let x vary continuously from 1 to the point defined by $-e^{-\alpha-\pi i} = e^{-\alpha x}$ so that $w = J(-e^{-\alpha x})$ describes t. $w = 1/J(-e^{-\alpha x}) = J(-e^{-\alpha/x})$ will then describe a path t' in the lower half plane, likewise starting from $x = 1$ and ending at a point given by $-e^{-\alpha/2+\pi i/2} = -e^{-\alpha/x}$, since $-ie^{-\alpha/2} = 2 = 1/J(-e^{-\alpha-\pi i})$ and $J(z) = 2$ has no other solution except $z = -ie^{-\alpha/2}$ inside the half circle $\mathrm{Im}\{z\} \leq 0$, $|z| \leq e^{-\alpha/2}$. (This will be shown in detail in the proof of Theorem V, Lemma II.) The equation

$$J(-e^{-\alpha-\pi i}) \cdot J(-e^{-\alpha/2+\pi i/2}) = 1$$

therefore entails the existence of a number x_0 such that

$$-\alpha - \pi i = -\alpha x_0$$

and

$$-\alpha/2 + \pi i/2 = -\alpha/x_0,$$

whence

$$(\alpha + \pi i)(\alpha - \pi i) = 2\alpha^2,$$

$$\alpha^2 + \pi^2 = 2\alpha^2,$$

$$\alpha^2 = \pi^2.$$

α being positive, we have therefore

$$\alpha = \pi.$$

5. Relation (6) also contains the fact that $J(e^{\pi i \tau})$, $(J_m\{\tau\} > 0)$, is an automorphic function of τ. For if we write—as usual in the theory of elliptic functions—$k^2(\tau) = J(e^{\pi i \tau})$, (6) takes the form

$$k^2(\tau) + k^2[-(1/\tau)] = 1. \tag{9}$$

Because of $J(e^{\pi i \tau}) = J(e^{\pi i (\tau \pm 2)})$ we have further

$$k^2(\tau \pm 2) = k^2(\tau). \tag{10}$$

By combining (9) and (10) we obtain

$$1 - k^2[-(1/\tau)] = k^2(\tau) = k^2(\tau \pm 2) = 1 - k^2[-(1/(\tau \pm 2))],$$

i. e.,

$$k^2[-(1/\tau] = k^2[-(1/(\tau \pm 2))].$$

On writing $-1/\tau$ instead of τ, this becomes

$$k^2(\tau) = k^2[\tau/(1 \pm 2\tau)].$$

Now it can be easily verified that by a suitable combination of the transformations $\tau \to \tau \pm 2$ and $\tau \to \tau/(1 \pm 2\tau)$ any transformation $\tau \to (\alpha\tau + \beta)/(\gamma\tau + \delta)$ can be obtained, where α and δ are two odd, and β and γ two even integers satisfying $\alpha\delta - \beta\gamma = 1$. For such integers α, β, γ, δ we have, therefore,

$$k^2[(\alpha\tau + \beta)/(\gamma\tau + \delta)] = k^2(\tau).$$

PART II.

1. We now begin a discussion of a class of analytic functions, first considered by Hurwitz, which we shall denote by H. A function $w = f(z) = a_1 z + a_2 z^2 + \cdots$ is said to belong to H if $f(z)$ is regular in the unit circle and does not vanish there except for $z = 0$.

The elliptic modular function $J(z)$ clearly belongs to H; moreover, any function $w = f(z)$ of H which, for $|z| < 1$, omits a finite value $w = d$, can be represented with the help of $J(z)$ and a bounded function $\omega(z)$. Indeed, if we consider the function

$$\omega(z) = J^{-1}[f(z)/d],$$

we see that, starting from $z = 0$, $\omega(z)$ may be uniformly continued within the whole circle $|z| < 1$, since the critical values of $J^{-1}[w]$ $(w = 0, 1, \infty)$ are not taken by $w = f(z)/d$ for $0 < |z| < 1$. Besides, $|J^{-1}[w]| < 1$ for all values of w differing from $0, 1, \infty$. Hence,

$$f(z) = d \cdot J[\omega(z)], \tag{13}$$

where $\omega(z)$ is regular for $|z| < 1$ and

$$|\omega(z)| \leq |z|, \tag{13a}$$

i. e., $f(z)$ is subordinate to $d \cdot J(z)$.

From (13) and (13a), many properties of the functions $f(z)$ belonging to H may be inferred. The first result to be mentioned here is the following classical result due to Hurwitz:

THEOREM I. *A function* $w = f(z) = z + a_2 z^2 + \cdots$ *belonging to* H *takes, for* $|z| < 1$, *every value inside the circle*

$$|w| = 1/16.$$

The constant $1/16$ *is the best possible.*

The proof of this theorem follows immediately from the representation (13), by which

$$1 = d \cdot \omega'(0) \cdot J'(0),$$

where d was a value not taken by $f(z)$ for $|z| < 1$. Remembering that $J'(0) = 16$, we have therefore $1 \leq 16\,|d|$, i. e., $|d| \geq 1/16$.

Remark. It is worthy of note that in the representation $f(z) = d \cdot J[\omega(z)]$, $\omega(z)$ is not only bounded but satisfies also $\omega(z) \neq 0$ for $z \neq 0$; on the other hand, in the proof of Theorem I—as also in the proofs of other theorems to follow—only the fact that $\omega(z)$ is bounded is made use of. Consequently, these theorems also remain true for the more general class of functions $g(z) = d \cdot J[\omega_1(z)]$, where $\omega_1(z)$ is an otherwise unrestricted bounded function. This class of functions, say H_d, may also be characterized as follows: A function $w = g(z) = b_1 z + b_2 z^2 + \cdots$ belongs to H_d if $g(z)$ is regular and $g(z) \neq d$ for $|z| < 1$ and if any zero $z = \alpha$ of $g(z)$ may be connected with $z = 0$ by a curve $z = z(t)$ such that the curve $w(t) = w[z(t)]$ does not surround the point $w = d$.

We therefore arrive at the following generalization of Theorem I:

THEOREM Ia. *Let* $w = g(z) = z + b_2 z^2 + \cdots$ *be a function regular for* $|z| < 1$; *then any value* $w = d$ *not taken there by* $w = g(z)$ *and not surrounded by a curve* $w(t) = g[z(t)]$, *where the curve* $z(t)$ *connects* $z = 0$ *with another zero of* $g(z)$, *satisfies*

$$|d| \geq 1/16.$$

2. An immediate consequence of (13) is

$$\underset{|z|=\rho}{|f(z)|} \leq |d| \underset{|z|=\rho}{\operatorname{Max}} |J(z)|.$$

Although the best possible, this inequality is not very convenient to handle. A more convenient bound for $|f(z)|$ —which, of course, is not "the best possible" any more—is given by

THEOREM II. *If $f(z)$ belongs to H and $f(z) \neq d$ for $|z| < 1$, then*

$$|f(z)| \leq (|d|/16)e^{-\pi^2/\log\rho}, \qquad [|z| = \rho].$$

Proof. From the product expansion (4) which, by replacing z by $-z$, takes the form

$$-J(-z) = 16z\prod_{n=1}^{\infty}[(1+z^{2n})/(1-z^{2n-1})]^8, \tag{14}$$

we may conclude that all the coefficients of the expansion of $-J(-z)$ in powers of z are non-negative, this being true of each single factor $[(1+z^{2n})/(1-z^{2n-1})]$ making up this product. This entails the equality

$$\operatorname*{Max}_{|z|=\rho}|J(z)| = -J(-\rho)$$

and the inequality

$$-J(-\rho) \geq 16\rho.$$

In view of formula (7) which, replacing $e^{-\pi}$ by ρ, may be written

$$J(-\rho)J(-e^{\pi^2/\log\rho}) = 1,$$

we, therefore, obtain

$$\begin{aligned}|f(z)| &\leq |d| \operatorname*{Max}_{|z|=\rho}|J(z)| = -|d|J(-\rho)\\ &= -(|d|)/[J(-e^{-\pi^2/\log\rho})] \leq (|d|/16)\,e^{-\pi^2/\log\rho}.\end{aligned}$$

It should be noted that this inequality gives the correct order of growth of $\operatorname*{Max}_{f(z)\subset H}|f(z)|$ for $\rho\to 1$, since $\lim_{\rho\to 0} -(J[-\rho]/16\rho) = 1$ entails $\lim_{\rho\to 1} -J(-\rho)\cdot 16e^{\pi^2/\log\rho} = 1$.

3. By Theorem I, the modulus of the first coefficient of the Taylor expansion of $f(z)$ is smaller than or equal to the first coefficient of the expansion of $-|d|J(-z)$. We shall now show that the same inequality holds between all other corresponding coefficients of $f(z)$ and $-|d|J(-z)$.

THEOREM III. *Let $f(z) = \sum_{n=1}^{\infty} a_n z^n$ belong to H and let d be a value such that $f(z) \neq d$ for $|z| < 1$; then*

$$|a_n| \leq |d| A_n,$$

where A_n $(n = 1, 2, \cdots)$ *are the coefficients of the expansion*

$$-J(-z) = \sum_{n=1}^{\infty} A_n z^n.$$

Proof. As already seen, all coefficients A_n are non-negative. Furthermore, if we denote the first and second differences $A_n - A_{n-1}$ and $A_n - 2A_{n-1} + A_{n-2}$ by $D_n^{(1)}$ and $D_n^{(2)}$ respectively, we have, by (14),

$$-(1-z)J(-z) = \sum_{n=1}^{\infty} D_n^{(1)} z^n = 16z \frac{(1+z^2)^8}{(1-z)^7} \prod_{n=2}^{\infty} \left(\frac{1+z^{2n}}{1-z^{2n-1}}\right)^8$$

and

$$-(1-z)^2 J(-z) = \sum_{n=1}^{\infty} D_n^{(2)} z^n = 16z \frac{(1+z^2)^8}{(1-z)^7} \prod_{n=2}^{\infty} \left(\frac{1+z^{2n}}{1-z^{2n-1}}\right)^8$$

(with $D_1^{(1)} = A_1$, $D_1^{(2)} = A_1$, $D_2^{(2)} = A_2 - 2A_1$). Both products being composed of factors the expansion of which in powers of z have non-negative coefficients, all numbers $D_n^{(1)}$ and $D_n^{(2)}$ are non-negative. Now, by a well known theorem,[1] the coefficients a_n of a function $f(z) = \sum_{n=1}^{\infty} a_n z^n$ subordinate to another function $F(z) = \sum_{n=1}^{\infty} a'_n z^n$ with $a'_1 > 0$, $a'_2 - a'_1 \geq 0$, $a'_n - 2a'_{n-1} + a'_{n-2} \geq 0$ $(n \geq 3)$, satisfy the inequality $|a_n| \leq a'_n$. This completes the proof.

Although Theorem III gives the exact bounds for the coefficients a_n, it is sometimes more desirable to have a convenient, though inexact, bound for the coefficients a_n, since the computation of the coefficients A_n becomes rather tedious for large n. Such a bound is given by

THEOREM IV. *If* $f(z) = \sum_{n=1}^{\infty} a_n z^n$ *belongs to* H *and* $f(z) \neq d$ *for* $|z| < 1$, *we have*

$$|a| \leq (|d|/16) e^{2\pi\sqrt{n}}.$$

This is an immediate consequence of Theorem II. Indeed:

$$|a_n| \leq |1/\rho^n| \operatorname*{Max}_{|z|=\rho} |f(z)| \leq (|d|/16)\rho^{-n} \cdot e^{-(\pi^2/\log\rho)}.$$

[1] See J. E. Littlewood, *Lectures on the Theory of Functions*, Oxford University Press, 1944, p. 169.

For $\rho = e^{-(\pi/\sqrt{n})}$ this becomes

$$|a_n| \leq (|d|/16)e^{2\pi\sqrt{n}}.$$

4. We shall now prove the following theorem:

THEOREM V. *If $f(z) = z + a_2z^2 + \cdots$ belongs to H and $f(z) \neq \alpha$ for $|z| < 1$, then there exists a positive number $\rho = \rho(|\alpha|)$ such that $f(z)$ is univalent inside the circle $|z| = \rho$. For $|d| \geq e^{\beta}/16$, where β is the positive solution of*

$$\left(\frac{\sqrt{\beta+2}-\sqrt{\beta}}{\sqrt{\beta+2}+\sqrt{\beta}}\right)^{e^{-(\beta\sqrt{\beta}/[\beta+2])}} = e^{-(\pi/2)},$$

we may take

$$\rho = 1 + \log 16\,|d| - \sqrt{(1+\log 16\,|d|)^2 - 1},$$

and for $|\alpha| < e^{\beta}/16$, the positive solution of

$$\rho \cdot [16\,|d|]^{(\rho-1)/(\rho+1)} = e^{-(\pi/2)}.$$

The value of ρ given for the case $|d| \geq e^{\beta}/16$ is the best possible, while in the case $|d| e^{\beta}/16$ the value given tends to the best possible only for $|d| \to 1/16$.

For the proof of this theorem we require the following two lemmas:

LEMMA I. *Let $\omega(z) = \alpha z + c_2z^2 + \cdots$ be regular and $|\omega(z)| \leq |z|$ for $|z| < 1$, and let $\omega(z)$ further satisfy $\omega(z) \neq 0$ for $0 < |z| < 1$; then $\omega(z)$ is univalent inside the circle $|z| = \rho$ with*

$$\rho = 1 + \log 1/|\alpha| - \sqrt{(1+\log 1/|\alpha|)^2 - 1}.$$

This value of ρ is the best possible.

Proof. By hypothesis, $\omega(z)/z \neq 0$ and $|\omega(z)/z| \leq 1$ for $|z| < 1$. Hence, the function

$$g(z) = \log [z/\omega(z)] = \log 1/\alpha - (c_2/\alpha)z + \cdots$$

is regular for $|z| < 1$ and we have there $R\{g(z)\} > 0$. By a classical result, this entails

$$R\{g(0)\}[(1-|z|)/(1+|z|)] \leq R\{g(z)\} \leq R\{g(0)\}[(1+|z|)/(1-|z|)] \tag{15}$$

and

$$|g'(0)| \leq 2R\{g(0)\}. \tag{16}$$

In view of (15), we have

$$\log |z/\omega(z)| \geq \log 1/|\alpha|[(1-|z|)/(1+|z|)],$$

i. e.,

$$|\omega(z)| \leq |z| \cdot |\alpha|^{(1-|z|)/(1+|z|)}. \tag{17}$$

Since the linear substitution $z^* = (x-z)/(1-\bar{x}z)$ ($|x| < 1$) transforms the unit circle into itself, we have also $R\{g(z^*)\} > 0$. Now the first terms of the expansion of $h(z) = g(z^*)$ in powers of z are

$$\begin{aligned} h(z) &= g[(x-z)/(1-\bar{x}z)] = g(x) - (1-|x|^2)g'(x)z + \cdots \\ &= \log x/\omega(x) - \{(1-|x|^2)/x\}\{[1-(x\omega'(x))/\omega(x)]\}z + \cdots. \end{aligned}$$

The real part of $h(x)$ being positive for $|z| < 1$, we have, by (16),

$$|h'(0)| \leq 2R\{h(0)\},$$

whence

$$\{(1-|x|^2)/|x|\}\ |1-(x\omega'(x)/\omega(x))| \leq 2\log|x/\omega(x)|,$$

i. e.,

$$|1-(x\omega'(x)/\omega(x))| \leq 2\{|x|/(1-|x|^2)\}\log|x/\omega(x)|. \tag{18}$$

The necessary and sufficient condition for the circle $|z| < \rho$ to be mapped by $w = \omega(z)$ on a schlicht and star-shaped domain is $R\{z\omega'(z)/\omega(z)\} \geq 0$ for $|z| < \rho$. This condition is certainly fulfilled if $|1-(z\omega'(z)/\omega(z))| \leq 1$. By (18), this inequality holds for values z satisfying

$$\{2|z|/(1-|z|^2)\}\log|z/\omega(z)| \leq 1.$$

By (15), we have

$$\log|z/\omega(z)| \leq \log\{1/|\alpha|\}[(1+|z|)/(1-|z|)],$$

i. e.,

$$\{2|z|/(1-|z|^2)\}\log|z/\omega(z)| \leq \log 1/|\alpha| \cdot 2|z|/(1-|z|)^2.$$

$\omega(z)$ will therefore be univalent for $|z| < \rho$ with ρ satisfying

$$\log 1/|\alpha| \cdot 2\rho/(1-\rho)^2 = 1$$

i. e.,

$$\rho = 1 + \log 1/|\alpha| - \sqrt{(1+\log 1/|\alpha|)^2 - 1}.$$

That this result is the best possible is shown by the function

$$\omega_0(z) = z \cdot e^{\log|\alpha|[(1+z)/(1-z)]} = |\alpha| z + \cdots$$

for which $|\omega(z)| \leqq |z|$ and $\omega(z) \neq 0$ for $0 < |z| < 1$, as required by the hypotheses of the lemma. As easily verified, the derivative of $\omega_0(z)$ vanishes at $z = \rho$ and $\omega_0(z)$ is therefore certainly not univalent in a circle $|z| < r$ with $r > \rho$.

LEMMA II. *The radius of univalency of the elliptic modular function* $J(z)$ *is* $\rho = e^{-(\pi/2)}$.

Proof. By (1), $J(z_1) = J(z_2)$ is equivalent to

$$4\sqrt{J(z_1^2)}/[1 + \sqrt{J(z_1^2)}]^2 = 4\sqrt{J(z_2^2)}/[1 + \sqrt{J(z_2^2)}]^2.$$

This equation has two solutions, viz., $J(z_1^2) = J(z_2^2)$ and $J(z_1^2) \cdot J(z_2^2) = 1$. If z_1 and z_2 are both assumed to lie on the circumference of the circle of univalency of $J(z)$, $J(z_1^2) = J(z_2^2)$ is clearly impossible and we must necessarily have $J(z_1^2) \cdot J(z_2^2) = 1$. As shown before, $|J(z)|$ attains its maximum on the circle $|z| = r$ for $z = -r$; we further had $J(e^{-\pi}) = \frac{1}{2}$ and $J(-z) = J(z)/(J(z) - 1)$, whence $J(-e^{-\pi}) = -1$. For values z satisfying $|z^2| < e^{-\pi}$ we therefore have $|J(z^2)| < 1$. Consequently, if $|z_1| < e^{-(\pi/2)}$ and $|z_2| < e^{-(\pi/2)}$, we certainly have $J(z_1^2) \cdot J(z_2^2) \neq 1$, which is equivalent to $J(z_1) \neq J(z_2)$. Therefore, $J(z)$ is univalent in the circle $|z| < e^{-(\pi/2)}$. On the other hand, in virtue of (1):

$$J(\pm ie^{-(\pi/2)}) = 4\sqrt{(-e^{-\pi})}/[1 + \sqrt{J(-e^{-\pi})}]^2 = \pm 4i/(1 \pm i)^2 = 2,$$

i. e., both $J(ie^{-(\pi/2)})$ and $J(-ie^{-(\pi/2)})$ have the value 2. Hence, $\rho = e^{-(\pi/2)}$ is the exact radius of univalency of $J(z)$.

In order now to prove Theorem V itself, we note that the univalency of $f(z) = d \cdot J[\omega(z)]$, $(|\omega(z)| \leq |z|)$, can break down owing to two different reasons. These are: 1) $\omega(z)$ ceases to be univalent; 2) although $\omega(z_1) \neq \omega(z_2)$, we have $J[\omega(z_1)] = J[\omega(z_2)]$. Since, by Lemma II, the case $J[\omega(z_1)] = J[\omega(z_2)]$, $\omega(z_1) \neq \omega(z_2)$ cannot happen for $|\omega(z_1)| < e^{-(\pi/2)}$ and $|\omega(z_2)| < e^{-(\pi/2)}$ and, by Lemma I, $\omega(z) = z/16d + \cdots$ is univalent for $|z| < \rho(|d|) = 1 + \log 16\,|d| - \sqrt{(1 + \log 16\,|d|)^2 - 1}$, $f(z)$ will also be univalent for $|z| < \rho(|d|)$. By (17), $|\omega(z)| \leq |z|\,[16\,|d|]^{|z|-1/|z|+1}$. Accordingly, the formula for the radius of univalence derived just now will

hold good as long as $|d|$ is larger than the positive solution $|d_0|$ of the equation obtained by eliminating ρ from the equations

$$\rho = 1 + \log 16\,|d_0| - \sqrt{(1+\log 16\,|d_0|)^2 - 1}$$

and

$$\rho \cdot [16\,|d|]^{\rho-1/\rho+1} = e^{-(\pi/2)},$$

i. e., $|d_0| = e^{\beta}/16$, where β is the positive solution of

$$\left(\frac{\sqrt{2+\beta}-\sqrt{\beta}}{\sqrt{2+\beta}+\sqrt{\beta}}\right) e^{-\beta\sqrt{\beta/(2+\beta)}} = e^{-(\pi/2)}.$$

If $|d|$ is smaller than $|d_0|$, the formula for the radius of univalency of $\omega(z)$ of course still holds, but since $|\omega(z)|$ may now be larger than $e^{-(\pi/2)}$, the multivalency of $J(z)$ may come into play. This can be avoided by choosing $|z|$ smaller than the positive solution ρ of

$$\rho[16\,|d|]^{(\rho-1)/(\rho+1)} = e^{-(\pi/2)}$$

Since, for $|d| < |d_0|$, this value ρ is smaller than the radius of univalency of the functions $\omega(z)$, $f(z) = d \cdot J[\omega(z)]$ will therefore be univalent for $|z| < \rho$. This completes the proof.

THE HEBREW UNIVERSITY, JERUSALEM.

ON THE LAPLACE-FOURIER TRANSCENDENTS OCCURRING IN MATHEMATICAL PHYSICS.*

By AUREL WINTNER.

---

1. The success of Laplace's method of adjoints in case of a given differential equation (which is ordinary, linear, homogeneous and will be assumed to be of second order) always depends on the discovery of a "suitable" path of integration. The existence or non-existence of such paths is sensitive under arbitrarily small modifications of the coefficient functions. The situation is made particularly unstable by the implied demands of what are called the applications, which (for reasons of reality) prefer *rectilinear* paths and, at the same time, *positive* density functions (as to this terminology, cf. [6], pp. 226-240, 274-285). One could almost say that the relevant "classical transcendents of mathematical physics" have been added to the list of "known functions" because their differential equations present some of the few cases in which a contour and the method of contour deformations have, accidentally, succeeded in the above sense. Correspondingly, the only essential idea which, since the eighteenth century, has been added to the subject is Cauchy's theory of contour integrations, that is, his justification of Laplace's indiscriminate use of complex variables.

Thus it is natural to desire a replacement of the substantially finite set of classical examples by a theoretical approach which, upon an inspection of the differential equation, should be able to guarantee the existence of a solution representable as a definite integral of the type in question. Needless to say, such a plan cannot succeed if it involves, as usual, boundary conditions and contour deformations, the issue being precisely the existence or non-existence of an appropriate path of integration. Inasmuch as the desideratum is surely not novel (it could be as old as the method itself), it is somewhat unexpected that, as will be shown below, it can be realized, and in quite a "practicable" or "explicit" fashion, in *real* terms which are available since a long time (though not of course since Cauchy's time).

All that will be needed is a combination of two theories, both of which concern the real domain. The first of them is contained in A. Kneser's elementary considerations (1896) on the characteristics of differential equa-

---

* Received September 12, 1946.

tions of second order, made easily accessible by Borel's lectures on divergent series (1901). The second, available only since about a quarter of a century, is supplied by the existence theorem of Hausdorff-Bernstein on complete monotony, a result generally known. Correspondingly, what is somewhat unexpected is, not that the latter theorem, originally found in its own pure context, turns out to supply an answer to an old "applied" question, but the circumstance that this does not seem to have been observed before.

2. The above formulation of the desideratum is such as to make it clear that the criterion a priori cannot dispose of the necessity of explicit calculations, if explicit results are wanted. In this regard, the resulting situation can be described as follows: In the classical approach, there are two unknown elements; namely, a suitable integration path, which need not exist and must be guessed when it does, and a corresponding density function (the "lower function" of the transformation), which, if it exists, is supplied by a suitable solution of the adjoint differential equation. In contrast, the criterion a priori will assure the existence of a suitable integration path and of a corresponding density function. In addition, it will supply the determination of a suitable integration path, which will simply be the half-line $w = u$, $(0 < u < \infty)$, in Laplace's case and the half-line $w = iv$, $(0 < v < \infty)$, in Fourier's case. Moreover, the density function will be guaranteed to be real and non-negative. Since the possibility of its vanishing (though not, of course, of its identical vanishing) is included in its being non-negative, this will comprise the case of integration paths which are finite segments, or half-lines terminating at a point $z = c \neq 0$. But what will not (and, by the very nature of the desideratum, should not) be furnished is an explicit determination of that non-negative density function.

The calculation leading to such a determination must therefore remain dependent on a solution of the adjoint of the given differential equation or on equivalent means. Actually, the given differential equation being of second order, it will (at least in principle) be admissible to assume it in the normalized self-adjoint form,

$$x'' + f(t)x = 0. \tag{1}$$

The last parenthetical proviso is necessary, since any relevant criterion must surely be so sensitive as to fail to be covariant under transformations leading to the normal form (1); the more so as the coefficient function, say $g = g(t)$, which occurs in the general form, $(gx')'$, of a self-adjoint differential operation of second order is made to be $g \equiv 1$ (cf. Liouville's transformation). But the

normal form (1) suffices for the presentation of all the essential points in the procedure.

3. A systematic embedding of classical transcendents into the general theory will be developed elsewhere. The following remarks, illustrating the theory in the simplest case, that of Bessel's equation,

$$y'' + y'/r + (1 - \alpha^2/r^2)y = 0,$$

can therefore be somewhat high-handed.[1] In this case, one of the substitutions leading to the normal form (1) is simple indeed:

$$r = t, \qquad y = t^{-\frac{1}{2}}x;$$

the coefficient of (1) then becomes

(1 bis) $$f(t) = 1 - \beta/t^2, \text{ where } \beta = (\alpha + \tfrac{1}{2})(\alpha - \tfrac{1}{2})$$

(for another, function-theoretically superior, substitution and the corresponding $f(t)$, cf. [4], p. 157). The index, α, is here arbitrary. However, the classical "real" representation of $J_\alpha(r)$ as ($r^{-\alpha}$ times) a Fourier transform is valid only when $\alpha > -\frac{1}{2}$. Instead of this, suppose that $\beta > 0$. Then, by the last formula line, $f(t)$ becomes the simplest of those polynomials in $1/t^2$ which have real coefficients with alternating signs (and with a positive zeroth coefficient).

It turns out that this property alone is sufficient for the existence of a solution of (1) which is a (generalized) Fourier transform of a non-negative density function. Moreover, nothing is changed if the polynomial is replaced by a transcendental entire function. In addition, even the requirement of an alternating coefficient sequence is unnecessarily strict. In fact, what it requires is that the coefficients should become positive after the substitution $t \to it$. But this is sufficient in order that $f(it)$ be completely monotone (in fact, the n-th derivatives of the functions $1/t^2$, $1/t^4$, $1/t^6, \cdots$, where $0 < t < \infty$, are positive or negative according as n is even or odd). However, all that really matters is the complete monotony of the even function $f(it)$ on the half-line $t > 0$:

---

[1] What will be disregarded is the necessity of replacing the relevant solution, $y = J_\alpha(r)$, by the corresponding Fourier integral, which is $r^{-\alpha}$ times $J_\alpha(r)$; the additional transition to (1), which modifies the factor $r^{-\alpha}$ to $r^{-\alpha-\frac{1}{2}}$; finally, the determination of the "correct branch" when r is replaced by ri in this factor (the exponent, $-\alpha - \frac{1}{2}$, is an integer only in the elementary cases of J_α).

If $f(1/z)$ is an even entire function in z and if

$$(2) \qquad (-1)^n d^n f(it)/dt^n \geqq \text{when } 0 < t < \infty,$$

where $n = 0, 1, \cdots$, (for instance, if

$$f(z) = \sum_{n=0}^{\infty} (-1)^n a_n / z^{2n} \text{ when } z \neq 0,$$

where $a_n \geqq 0$), then the differential equation (1) possesses a Fourier solution $x(\not\equiv 0)$, of the form

$$(3) \qquad x(t) = \left(\int_0^{\infty}\right) e^{itr} d\phi(r) \text{ if } 0 < t < \infty,$$

where ϕ ($\neq$ const.) is a monotone, but not necessarily bounded, function on the half-line $0 \leqq r < \infty$.

Since this half-line is closed, the possible unboundedness of ϕ takes place, of course, only when $r \to \infty$. The parenthetical integral sign in (3) refers to the Abelian value of the Fourier-Stieltjes transform. By this is meant that, if t (> 0) is fixed, the integral

$$\int_0^{\infty} e^{-\epsilon r} d\phi_t(r), \text{ where } d\phi_t(r) = e^{itr} d\phi(r),$$

is convergent when $\epsilon > 0$ and tends, as $\epsilon \to 0$, to a finite limit, which is (3) (the existence of this limit is part of the assertion). According to the integral form of Abel's lemma on power series, (3) is identical with

$$\int_0^{\infty} e^{itr} d\phi(r), \qquad (t > 0),$$

provided that the latter integral (which, since the total variation of $\phi(r)$ can be ∞, must be meant as an improper Stieltjes, rather than as a Lebesgue-Stieltjes, integral) is convergent; in which case (3) is sure to exist. Incidentally, it remains undecided whether the replacement of the last formula line by (3) is actually necessary (for *some* f satisfying the assumptions of the theorem).

The assumptions required of f are satisfied if $f(t) \equiv 1$ (this is the limiting case, $\beta = 0$, of the normalized Bessel equation considered above). Then (1) becomes the differential equation $x'' + x = 0$, having the linearly

independent solutions $x = \cos t$, $x = \sin t$ which, since $f(t)$ is real, are equivalent to the single solution $x = e^{it}$. Inasmuch as the latter results by choosing $\phi(r)$ in (3) to be the step function $\frac{1}{2}\operatorname{sgn}(r-1)$, the appearance of a Stieltjes integral in this quite analytical context becomes understandable to some extent.

4. The function $\phi(r)$ must be absolutely continuous, in fact, piecewise analytic, if it can be obtained by the classical method of the adjoint differential equation. But, as exemplified by the function $J_0(t)$, a Fourier transform (3) can (even though, as shown by the example of $t^{-\frac{1}{2}}$ or $\exp(-t^2)$, need not) perform an infinity of "oscillations" or "waves," when t varies from 0 to ∞. On the other hand, there are no *formal techniques* leading, by some method of comparing coefficients, to "*wave*" *solutions* of the type (3) (with $d\phi \geqq 0$), when $f(t)$ in (1) is a given function. This is precisely the reason why all direct attempts trying to satisfy a (suitable, but still unspecified) differential equation (1) by anything like (3) could not succeed. However, the transformation $t \to it$, used above, replaces each of the partial vibrations, e^{irt}, of the bundle (3) by the corresponding damping, e^{-rt}, and the corresponding bundle cannot, of course, have "waves."

Actually, the solution curve which thus results upon the transformation $t \to it$ becomes so *smooth in the large* that a combination of the results of Kneser and of Hausdorff-Bernstein, referred to in the introduction, becomes applicable. Correspondingly, the criterion italicized above will not have to be considered directly, since it results as a straightforward corollary of the following theorem:

If a function $f(t)$, defined on the open half-line $t > 0$, has derivatives of arbitrarily high order satisfying

$$(4_n) \qquad (-1)^n f^{(n)}(t) \leqq 0, \qquad 0 < t < \infty,$$

where $n = 0, 1, \cdots$, then the differential equation (1) *possesses a Laplacean solution $x(\not\equiv 0)$, of the form*

$$(5) \qquad x(t) = \int_0^\infty e^{-tr}\, d\phi(r), \qquad 0 < t < \infty,$$

where ϕ ($\not\equiv$ const.) is a monotone, but not necessarily bounded, function on the closed half-line $r \geqq 0$.

That this latter theorem implies the former as a corollary, follows from the circumstance that, if t in (5) is replaced by a complex variable, say z, then, since the convergence of the integral (5) on the half-line $t > 0$ is part of the latter theorem, the resulting integral converges to a regular function in the half-plane $\Re z > 0$. Thus the former theorem can be obtained by letting $\Re z \to 0$ and taking into account the definition of the parenthetical integral (3). In fact, every solution of a linear differential equation (1) (in which t is now complex) is regular at every point t at which the coefficient function, $f(t)$, is regular.

5. The second theorem is of a more symmetric nature, since it can simply be expressed as follows: *If $-f$ is completely monotone, then* (1) *has a completely monotone solution* x ($\not\equiv 0$); $0 < t < \infty$. In fact, the existence theorem of Hausdorff-Bernstein states that (5), where $d\phi \geqq 0$, is always implied by (and, of course, implies) the conditions

$$(6_n) \qquad (-1)^n x^{(n)}(t) \geqq 0, \qquad 0 < t < \infty,$$

where $n = 0, 1, \cdots$. The formulation in terms of (4_n) and (5) is preferable only because, f being the given function, condition (4_n) can easier be checked than the existence of an integral representation, whereas what is wanted of the unknown solution is precisely the existence of an integral representation.

If $x(t)$ is a completely monotone solution ($\not\equiv 0$) of (1), another such solution is $Cx(t)$, where C is an arbitrary positive constant (this agrees with (5), where ϕ can be replaced by $C\phi$). But *only* this C remains undetermined: *Two completely monotone solutions of* (1) *are linearly dependent.* In fact, (5) implies that $x(t) = O(1)$ as $t \to \infty$. It does not imply that $x(t) = o(1)$, since $\phi(r)$ can have a jump at $r = 0$ (as illustrated by the simplest case, $f(t) \equiv 0$, which, according to (1), (5), belongs to $x(t) \equiv 0$, $\phi(r) = \operatorname{sgn} r$). But it is clear that this effect of a possible jump of $\phi(r)$ at $r = 0$ is removed if $d\phi(r)$ is multiplied by r. Since the resulting integral is

$$\int_0^\infty e^{-tr} r\, d\phi(r) = -x'(t),$$

by (5), it follows that $x'(t) = o(1)$ as $t \to \infty$. Hence, if $x = x_1(t)$ and $x = x_2(t)$ are completely monotone, then since $x_j(t) = O(1)$ and $x_j'(t) = o(1)$ as $t \to \infty$, the Wronskian, $x_1 x_2' - x_2 x_1'$, is $o(1)$. However, the Wronskian of two solutions of (1) is always a constant and cannot, therefore, be $o(1)$ unless it vanishes identically. This proves the last italicized remark.

6. The theorem which remains to be proved is equivalent to the statement that, if f satisfies the infinity of conditions $(4_0), (4_1), \cdots$, then (1) has a solution $x(\not\equiv 0)$ satisfying $(6_0), (6_1), \cdots$. The proof will succeed only because the statement can be "finitized." In fact, it can be sharpened to the following assertion, in which m denotes a *fixed* non-negative integer:

(7_m) *If a function $f(t)$, where $0 < t < \infty$, is of class $C^{(m)}$ and satisfies the $m+1$ conditions $(4_0), \cdots, (4_m)$, then the differential equation* (1) *has a solution $x(\not\equiv 0)$ satisfying the $m+2$ conditions $(6_0), \cdots, (6_{m+1})$.*

This will be concluded, by complete induction, from the case $m = 0$, which is substantially equivalent to the following theorem:

(7*) *If $f(t)$ is a real-valued, non-positive, continuous function on the open half-line $0 < t < \infty$, then the differential equation* (1) *has a solution $x(t)$ which is non-negative and non-increasing at every point of the half-line (and does not vanish everywhere).*

Ad (7_0). The assumption of (7_0) is the same as that of (7*), namely, (4_0). The assertions of (7*) are (6_0) and (6_1), whereas (7_0) claims (6_2) also. However, since (1) means that

$$x'' = -fx, \tag{8_0}$$

(6_2) is implied by (6_0) and (7_0).

Ad (7_1). The assumptions of (7_1) are (4_1) (along with the existence of a continuous f') and the assumptions of (7_0). The new assertion is (6_3). However, from (8_0),

$$-x''' = fx' + f'x. \tag{8_1}$$

According to (4_0), (4_1) and (6_0), (6_1), both terms on the right of (8_1) are non-negative. Hence, (6_3) follows from (8_1).

Ad (7_2). The assumptions of (7_2) are (4_2) (along with the existence of a continuous f'') and the assumptions of (7_1). The new assertion is (6_4). However, from (8_1),

$$x'''' = -fx'' - 2f'x' - f''. \tag{8_2}$$

According to (4_0), (4_1), (4_2) and (6_0), (6_1), (6_2), all three terms on the right of (8_2) are non-negative. Hence, (6_4) follows from (8_2).

Since the binomial coefficients, introduced by m-fold differentiation of the product (8_0), are positive, it is clear that the play of alternating signs is preserved in the induction $m \to m+1$. Consequently, all that remains to be assured is (7*).

7. If the assumption, $f \leqq 0$, of (7*) is replaced by $f < 0$, then the assertion of (7*) can be refined as follows:

(7 bis) *If $g(t)$ is a positive, continuous function on the open half-line $0 < t < \infty$, then the differential equation*

$$x'' = g(t)x \tag{9}$$

has a solution $x = x(t)$ which is positive (and so, by (9), convex) and decreasing for $0 < t < \infty$.

Remark. *Neither*

$$x(+0) \neq \infty \tag{10}$$

nor

$$x(\infty) \neq 0 \tag{11}$$

is claimed in (7 bis).

Actually, (7*) can be concluded from (7 bis), if use is made of Helly's selection theorem and of the locally continuous dependence of the general solution on $f = -g$.

Incidentally, (7 bis) instead of (7*), along with the induction $m \to m+1$, is sufficient to assure the existence of a representation (5) in the case in which (4_n) is assumed for every n. In fact, $f(t)$ is regular at every t in this case, and so it is clear from the convexity assumption (4_2) that $f(t)$ cannot vanish at all unless it vanishes identically. But then the differential equation becomes $x'' = 0$ and possesses, therefore, the solution $x(t) \equiv 1$, which, of course, has the properties claimed.

Accordingly, only (7 bis) remains to be ascertained.

Let (7′) denote the sentence which becomes of the sentence (7 bis) if the open half-line, $0 < t < \infty$, is replaced by a closed half-line, say $1 \leqq t < \infty$, in the assumption and in the assertion of (7 bis).

What corresponds to the negation of (10) (namely, the possibility $x(1+0) = \infty$) cannot occur in the case of a closed half-line. The latter can therefore be attacked by an initial point, $t = 1$. Thus (7 bis) is somewhat more elaborate than (7′). However, (7 bis) can be reduced to (7′). In fact, if an arc, say $t_1 \leqq t \leqq t_2$, of a solution path, $x = x(t)$, is in the upper

half-plane ($x > 0$), then it is convex ("subharmonic"); if it is in the lower half-plane ($x < 0$), it is concave. This is clear from (9), since $g > 0$. Hence, if (7′) is granted, and if $x = x(t)$ denotes a solution supplied by (7′) on the half-line $1 \leqq t < \infty$, the extension of (7′) to (7 bis) is furnished by that solution of (9) on the interval $0 < t \leqq 1$ which is determined by the initial values $x(1)$, $x'(1)$.

This reduces everything to the truth of (7′). But (7′) is known. In fact, (7′) is contained in Kneser's geometrical considerations ([5], pp. 183-191), which deal with a non-linear generalization of (9), where $g > 0$. Correspondingly, Kneser's proof admits of simplifications in the linear case of (7′) (in this respect, cf. pp. 42-46 of Borel's monograph [2]; actually, the proof can be simplified still further).

8. For the sake of completeness, a final comment must be made; one concerning the impossibility of (11), which is claimed by Kneser ([5], p. 191) but is not claimed above. This discrepancy can be cleared up as follows:

Let $x_1(t), x_2(t), \cdots$ be a sequence of continuous (or, for that matter, regular analytic) functions on the half-line $1 \leqq t < \infty$, and let $t_1, t_2, \cdots$ be a monotone sequence of t-values satisfying $1 < t_n \to \infty$ as $n \to \infty$. Suppose that $x_n(t)$ is negative and concave ("superharmonic") on the half-line $t_n < t < \infty$; that $x_n(t)$ is positive, convex and decreasing on the interval $1 \leqq t < t_n$; finally, that all curves $x = x_n(t)$ meet when $t = 1$ (and, for the sake of simplicity, that, if n and m are distinct, $x_n(t) = x_m(t)$ holds only when $t = 1$). Then it is clear, for instance from Helly's theorem, that the sequence $x_1(t), x_2(t), \cdots$ contains a subsequence which, uniformly on every bounded t-interval, tends to a limit function, say $x(t)$; and that $x(t)$ is positive, convex and decreasing on the whole half-line. In particular, $x(t)$ tends to a finite, non-negative limit as $t \to \infty$.

What Kneser seems to consider to be evident ([5], p. 182) is that, since $t_n \to \infty$ as $n \to \infty$ and $x_n(t_n) = 0$, this non-negative limit, $x(\infty)$, cannot be positive. Actually, it is easy to draw a sequence of curves which satisfy all the above conditions but lead to a positive $x(\infty)$. In fact, the convergence, uniform on every bounded t-interval of the half-line, need not be uniform on the whole of the half-line.

This explains the discrepancy. It was mentioned because it is important in connection with (3), (5). In fact, the inclusion of discontinuous functions $\phi(r)$ was seen to be an essential point in the above theorems on integral representations. But (11) is precisely the condition for a solution (5) in which $\phi(r)$ has no jump at $r = 0$.

Since the unknown of the problem is the function $\phi(r)$, it is worth while to mention that a jump $\phi(+0) \neq \phi(0)$ in (5) can be excluded by a criterion involving only the coefficient function of (1). In fact, such a criterion is contained in the following remark (which, instead of (4_n), assumes only that $f(t) \leqq 0$):

If $f(t)$ and $x(t)$ satisfy the assumptions of (7), then (11) cannot be true unless the integral*

$$\int^{\infty} f(t)dt \text{ is convergent.} \tag{12}$$

For, if $f(t)$ and $x(t)$ satisfy the assumptions of (7*), then, since $x(t)$ is convex by virtue of (1), the derivative $x'(t)$ must tend to 0 as $t \to \infty$. Since (1) means that

$$x'(t) + \int^{t} f(u)x(u)du \tag{13}$$

is independent of t, it follows that

$$\int^{\infty} f(t)x(t)dt \text{ is convergent,} \tag{14}$$

whether (12) be satisfied or not. Hence, it is clear from $f(t) \leqq 0$ that (11) is impossible if (12) is assumed (the existence of $x(\infty)$ being implied by (7*) alone).

Appendix.

Under the assumptions which had to be made for the existence of Fourier and Laplace solutions, (3) and (5), the coefficient function, $f(t)$, of (1) must of course be non-negative and non-positive, respectively. Correspondingly, the simplest illustrations of the "elliptic" case and of the "hyperbolic" case are $f(t) = a^2$ with $x(t) = \cos at$ and $f(t) = -a^2$ with $x(t) = e^{-at}$, respectively, where a is a positive constant; cf. (3), (5) and (1). In the case of e^{-at}, the possibility (11) does not take place. In the case of $\cos at$, the (real) solutions $x(t)$ are oscillatory, that is to say such as to possess an infinity of zeros (of first order). The limiting "parabolic" case, $f(t) \equiv 0$, of (1) satisfies the assumptions of both (5) and (3). But since all solutions of $x'' = 0$ are linear, each of its solutions $x(t) \not\equiv 0$ presents the exceptional possibility (11) and none of them is oscillatory. Hence, if (12) is interpreted

as an assumption the negation of which carries (1) "too close" to $x'' = 0$ in the hyperbolic case, $f(t) \leqq 0$, of (7*), it is reasonable to expect that the same negation will prevent the existence of non-oscillatory solutions in the elliptic case, $f(t) \geqq 0$.

This heuristic dualization suggests a counterpart of the last italicized remark. It is easy to see that the resulting dual happens to be true:

If $f(t)$, where $0 < t < \infty$, is a real-valued, non-negative, continuous function, then, unless (12) *is satisfied, every real-valued solution $x(t) \not\equiv 0$ of* (1) *is oscillatory.*

This remark, though quite on the surface (and possibly well-known), is not clear from a direct "comparison" (Sturm) of (1) with $x'' + a^2 x = 0$. In fact, the negation of (12), that is, the assumption

$$\int^{\infty} f(t)dt = \infty, \text{ where } f(t) \geqq 0, \tag{15}$$

is compatible with $\lim f(t) = 0$ (as is (12) with $\limsup |f(t)| = \infty$), as $t \to \infty$.

Since an interval containing two zeros of any (real-valued) solution $x(t) \not\equiv 0$ of (1) must contain a zero of any other such solution of (1) (Sturm), and since $x(t)$ can be replaced by $-x(t)$ in (1), it is sufficient to verify the following assertion: If (1) has a solution which is positive from a certain $t = t_0$ onward, then $f(t)$ cannot satisfy (15).

To this end, let the integral condition be disregarded. The remaining assumptions (namely, $f(t) \geqq 0$ and $x(t) > 0$, where $t > t_0$) imply, by (1), that $x''(t) \leqq 0$. In other words, the curve $x = x(t)$ is concave from below (from $t = t_0$ onward). Since it stays in the upper half-plane, it follows that $x(t)$ is non-negative and non-decreasing, whereas $x'(t)$ is non-negative and non-increasing. Hence, there exist a positive limit $x(\infty) \leqq \infty$ and a non-negative limit $x'(\infty) < \infty$. But (13) is independent of t by virtue of (1), and so the existence of a finite limit $x'(\infty)$ implies (14). Finally, (14) and the existence of a (finite or infinite) positive limit $x(\infty)$ necessitate (12), since $f(t) \geqq 0$.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES

[1] S. Bernstein, "Sur les fonctions absolument monotones," *Acta Mathematica*, vol. 52 (1928-29), pp. 1-66.

[2] E. Borel, *Leçons sur les séries divergentes*, Paris, 1901.

[3] F. Hausdorff, "Summationsmethoden und Momentfolgen, II," *Mathematische Zeitschrift*, vol. 9 (1921), pp. 280-299.

[4] F. Klein, *Vorlesungen über die hypergeometrische Funktion* (ed. O. Haupt), Berlin, 1933.

[5] A. Kneser, "Untersuchung und asymptotische Darstellung der Integrale gewisser Differentialgleichungen bei grossen reellen Werthen des Arguments, I," *Journal für die reine und angewandte Mathematik*, vol. 116 (1896), pp. 178-212.

[6] L. Schlesinger, *Einführung in die Theorie der gewöhnlichen Differentialgleichungen auf funktionentheoretischer Grundlage*, Berlin and Leipzig, 1922 (3rd ed.).

ON THE TAUBERIAN NATURE OF IKEHARA'S THEOREM.*

By AUREL WINTNER.

In case of ordinary Dirichlet series, Ikehara's theorem can be formulated as follows: If the series

$$f(s) = \sum_{n=1}^{\infty} a_n/n^s \tag{1}$$

is absolutely convergent in the half-plane $\sigma > 1$ and represents there a function which goes over into a continuous boundary function on the line $\sigma = 1$, then the Tauberian condition

$$a_n = O_L(1) \tag{2}$$

is sufficient to ensure that

$$\sum_{n=1}^{x} a_n = o(x). \tag{3}$$

This is not the usual formulation (cf. [6], pp. 127-130) but is readily seen to be equivalent to it. In fact, if a constant, c, is added to every a_n, then (2) remains unaltered, (1) goes over into $f(s) + c\zeta(s)$ and acquires therefore a "pole", of "residue" c, at $s = 1$, and, correspondingly, the $o(x)$ in (3) becomes $cx + o(x)$.

The proof of Ikehara's theorem depends on the same technique as Wiener's result on Lambert summability. The content of the latter result is (cf. [6], pp. 119-124) that the (L)-summability of a series $c_1 + c_2 + \cdots$ and the Tauberian condition

$$c_n/n = O_L(1) \tag{2 bis}$$

imply (A)-summability. The analogy is the more relevant as either of the theorems suffices for the transition from the non-vanishing of $\zeta(1 + it)$ to the prime number theorem. It is true that, in the Lambertian approach, recourse must be had to the Hardy-Littlewood lemma, according to which (A)-summability and (2 bis) together imply $(C, 1)$-summability (in fact, convergence as well; but this is not needed in the deduction of the prime number theorem), but Karamata [4] has discovered that this additional step

* Reecived September 12, 1946.

is simple indeed. For the general methodical situation, cf. [1], pp. 34-35, 46-47 and [3], pp. 38-40.

Actually, Wiener's Tauberian result on (L)-summability, when formulated in the above form, somewhat disguises the true situation. In fact, his result is just a corollary of an earlier theorem of Hardy and Littlewood [2], according to which (L)-summability *always* implies (A)-summability (whether a Tauberian restriction, such as (2 bis), be satisfied or not). But this theorem, though deeper than the prime number theorem itself, is not at all Tauberian in nature.

It is therefore natural to ask whether the situation is similar in the parallel case of Ikehara's approach to the prime number theorem; in other words, whether his Tauberian restriction, (2), is actually superfluous (at any rate, by virtue of the prime number theorem or, as the unconditional implication, $(L) \rightarrow (A)$, of Hardy and Littlewood, by virtue of a somewhat deeper theorem). The purpose of this note is *to answer this question in the negative.*

The situation would be simple indeed if, in the above formulation of Ikehara's theorem (which, as it stands above, is tautologous), it were not stipulated that the Dirichlet series be *absolutely* convergent in the half-plane $\sigma > 1$. For, if

$$a_n = (-1)^n n,$$

the Dirichlet series (1) is convergent in the half-plane $\sigma > 1$ and attains continuous boundary values on the line $\sigma = 1$; in fact, the function (1) becomes the entire function $(2^{2-s} - 1)\zeta(s-1)$. Nevertheless, since

$$\liminf_{x\to\infty} \sum_{n=1}^{x} (-1)^n n/x = -\tfrac{1}{2}, \qquad \limsup_{x\to\infty} \sum_{n=1}^{x} (-1)^n n/x = \tfrac{1}{2},$$

(3) is not true.

The construction of a Dirichlet series (1) having 1 as its abscissa of *absolute* convergence will be based on the general identity on which Riemann's proofs of the functional equation of $\zeta(s)$ depend. This identity, first justified by Perron [5], states that, if (1) is convergent in the half-plane $\sigma > 1$, then

$$f(s) = F(s)/\Gamma(s), \tag{4}$$

where

$$F(s) = \int_0^\infty x^{s-1} \sum_{n=1}^{\infty} a_n e^{-nx} dx, \qquad \sigma > 1. \tag{5}$$

Needless to say, the convergence of (1) for $\sigma > 1$ implies that the power series multiplying x^{s-1} in (5) is convergent for $e^{-x} < 1$, i. e., for every x in the *interior* of the integration domain (the integral (5) must therefore be interpreted as improper at both ends, $x = \infty$ and $x = 0$).

Suppose that

$$\int_0^{1-0} \left| \sum_{n=1}^{\infty} a_n r^n \right| dr < \infty. \tag{6}$$

Then the integral (5) is absolutely, hence uniformly, convergent in the closed strip $1 \leqq \sigma \leqq 2$. In fact, this is clear for the contribution,

$$\int_1^{\infty} x^{s-1} \sum_{n=1}^{\infty} a_n e^{-nx} dx.$$

of the integration range $1 \leqq x < \infty$. On the other hand, the contribution of the remaining range, $0 \leqq x \leqq 1$, to (5) is majorized, uniformly in the closed half-plane $\sigma \geqq 1$, by

$$\int_0^{\infty} \left| \sum_{n=1}^{\infty} a_n e^{-nx} \right| dx = \int_0^{1-0} \left| \sum_{n=1}^{\infty} a_n r^n \right| dr/r, \qquad (r = e^{-x}).$$

But the last integral differs from the integral (6) only in the factor $1/r$, which does not influence the convergence assumed for (6), since $1/r$ is bounded near $r = 1$, and is absorbed by r^n at $r = 0$, the summation index $n = 0$ being excluded.

Consequently, the function (5) is uniformly continuous (and, incidentally, bounded) in the open strip $1 < \sigma < 2$. Since $1/\Gamma(s)$ is not bounded in this strip, it does not follow that the function (4) is uniformly continuous there. But, since $\Gamma(1 + it) \neq 0$, it does follow that the function (4) represents in the half-plane $\sigma > 1$ a regular function which attains continuous boundary values on the line $\sigma = 1$.

Accordingly, it is sufficient to show that the following three conditions are compatible for a sequence $a_1, a_2, \cdots$:

(i) The Dirichlet series (1) is absolutely convergent when $\sigma > 1$.

(ii) Condition (6) is satisfied.

(iii) The estimate (3) fails to hold.

To this end, first choose an increasing sequence of positive integers $k_1, k_2, \cdots$ in such a way that, on the one hand,

$$\sum_{m=1}^{\infty} k_m^{-\epsilon} < \infty \text{ for every } \epsilon > 0 \tag{7}$$

and, on the other hand,

$$\sum_{m=1}^{\infty} k_m r^{k_m} = O(1-r)^{-3/2} \text{ as } r \to 1-0. \tag{8}$$

Such sequences exist (in fact, both conditions are satisfied by $k_m = m!^2$, the second, (8), being satisfied even by $k_m = m^2$).

It can also be assumed that the sequence $k_1, k_2, \cdots$ does not contain consecutive integers. Then a unique sequence $a_1, a_2, \cdots$ is defined by the following assignment: a_n is n, $-n+1$ or 0 according as n is a k_m, a k_m+1 or neither a k_m nor a k_m+1.

It is readily seen from this definition of a_n and from the first of the conditions, (7), imposed on the sequence $k_1, k_2, \cdots$, that (i) is satisfied. In order to ascertain (ii), it is sufficient to show that

$$\sum_{n=1}^{\infty} a_n r^n = O(1-r)^{-\frac{1}{2}} \text{ as } r \to 1-0. \tag{9}$$

But the definition of a_n means that the power series on the left of (9) is identical with

$$\sum_{m=1}^{\infty} k_m r^{k_m} - \sum_{m=1}^{\infty} k_m r^{k_m+1};$$

so that, since this difference can be contracted into

$$(1-r)\sum_{m=1}^{\infty} k_m r^{k_m},$$

(9) follows from (8).

Accordingly, only (iii) remains to be verified. However, by the definition of a_n,

$$\sum_{n=1}^{x} a_n = \sum_{k_m \leqq x} k_m - \sum_{k_m \leqq x-1} k_m. \tag{10}$$

Clearly, the difference on the right of (10) vanishes except at x-values

corresponding to its points of discontinuity, and the jumps amount to just the respective values of x itself. Accordingly,

$$\liminf_{x\to\infty} \sum_{n=1}^{x} a_n/x = 0 \text{ but } \limsup_{x\to\infty} \sum_{n=1}^{x} a_n/x = 1,$$

and so (3) fails to be true.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES

[1] G. H. Hardy, *Ramanujan*, Cambridge, 1940.

[2] G. H. Hardy and J. E. Littlewood, "On a Tauberian theorem for Lambert's series, and some fundamental theorems in the analytic theory of numbers," *Proceedings of the London Mathematical Society*, ser. 2, vol. 19 (1921), pp. 21-29.

[3] A. E. Ingham, *The Distribution of Prime Numbers*, Cambridge Tracts no. 30, 1932.

[4] J. Karamata, "Über die Hardy-Littlewoodsche Umkehrung des Abelschen Stetigkeitssatzes," *Mathematische Zeitschrift*, vol. 32 (1930), pp. 319-320.

[5] O. Perron, "Zur Theorie der Dirichletschen Reihen," *Journal für die reine und angewandte Mathematik*, vol. 134 (1908), pp. 95-143.

[6] N. Wiener, *The Fourier Integral and Certain of its Applications*, Cambridge, 1933.

UNION-PRESERVING TRANSFORMATIONS OF HIGHER ORDER SURFACE-ELEMENTS.*

By JOHN DE CICCO.

1. Curve-element transformations of the union-preserving type. In 1943, Kasner and De Cicco began the study of transformations in the plane and in space from differential elements of order n into those of first order.[1] We have termed such a correspondence *union-preserving* if it converts every union into a union. This class of union-preserving transformations includes the Lie theory of contact transformations as a special case.[2] In the previous work, our attention was devoted to the study of correspondences from curve-elements of order n into lineal-elements.[3] In the present article, we shall develop the theory of correspondences from surface-elements of order n into planar-elements.

For the purpose of contrasting our new theory with the previous study,[4] we shall state some of the results concerning union-preserving transformations in space from curve-elements of order n: $(x, y, z, y', z', \cdots, y^{(n)}, z^{(n)})$, where the accents denote total differentiation with respect to x, into lineal-elements: (X, Y, Z, Y', Z'). The general union-preserving transformations are determined by a single directrix equation of the form

$$\Omega(X, Y, Z, x, y, z, y', z', \cdots, y^{(n-2)}, z^{(n-2)}) = 0, \tag{C}$$

involving derivatives of order $(n-2)$, at most. If $n = 2$, this has the same form as the directrix equation of Lie. Although the directrix equations are the same for this case $n = 2$, the induced union-preserving transformation of second order is different in general from the induced contact transformation of Lie.

The special union-preserving transformations are obtained by extending

* Received August 15, 1946. Presented to the American Mathematical Society, 1946.

[1] Kasner and De Cicco, "Union-preserving transformations of differential elements," *Proceedings of the National Academy of Sciences*, vol. 29 (1943), pp. 271-275.

[2] Lie-Scheffers, *Beruhrungstransformationen*, Leipzig, 1896.

[3] Kasner and De Cicco, "A generalized theory of contact transformations," *Rivista de Matematicas de la Universidad Nacional de Tucuman*, Argentina, vol. 4 (1944), pp. 81-90.

[4] Kasner and De Cicco, "Union-preserving transformations of space," *Bulletin of the American Mathematical Society*, vol. 50 (1944), pp. 98-107. "Comparison of union-preserving and contact transformations," *Proceedings of the National Academy of Sciences*, vol. 32 (1946), pp. 152-156.

to lineal-elements transformations which carry curve-elements of order $(n-1)$ into points. Thus the class of union-preserving transformations from curve-elements of order n into lineal-elements consists of the general type defined by a single directrix equation of the form (C) or of the special type described above.

Any transformation not of the union-preserving type converts either precisely $\infty^{2(n+1)}$ or ∞^{∞} unions into unions.

The only available union-preserving transformations in the whole domain of curve-elements are firstly, the group of point transformations, and secondly, the set of union-preserving transformations from curve-elements or order n into lineal-elements, together with the extensions of these two types. This is an extension of Lie's theorem which states that the group of contact transformations of lineal-elements constitutes the extensions of the correspondences of the group of arbitrary point transformations.

2. Union-preserving transformations of higher order surface-elements. In this section, we shall state some of the results of the present article. It will be found that our new transformation theory is more analogous to the Lie theory than that described in the preceding section.

We shall study transformations in space from differential surface-elements of order n: $(x, y, z, p_{10}, p_{01}, \cdots, p_{j,m-j}, \cdots, p_{0n})$, where $p_{j,m-j} = \partial^m z/\partial x^j \partial y^{m-j}$ for $j = 0, 1, 2, \cdots, m$, and $m = 1, 2, \cdots, n$, into planar-elements (X, Y, Z, P, Q) where $P = \partial Z/\partial X$ and $Q = \partial Z/\partial Y$. Thus our transformations convert the $\infty^{(n^2+3n+6)/2}$ surface-elements of order n into the ∞^5 planar-elements. Our union-preserving transformations may be defined by a single directrix equation

$$(S) \qquad \Omega(X, Y, Z, x, y, z, p_{10}, p_{01}, \cdots, p_{j,m-j}, \cdots, p_{n-1,0}, \cdots, p_{0,n-1}) = 0,$$

involving partial derivatives of order $(n-1)$ at most; or by a pair or triplet of such directrix equations. For $n = 1$ this result coincides with the theorem of Lie concerning contact transformations of surface-elements of first order.

Thus our union-preserving transformations are defined by one arbitrary function of $(n^2+n+8)/2$ independent variables, or by two arbitrary functions of $(n^2+n+6)/2$ independent variables, or by three arbitrary functions of $(n^2+n+4)/2$ independent variables. Therefore in the notation introduced by Kasner, the contents of the three different classes of union-preserving transformations are $\infty^{1f[(n^2+n+8)/2]}$, $\infty^{2f[(n^2+n+6)/2]}$, and $\infty^{3f[(n^2+n+4)/2]}$, respectively.[5]

---

[5] Kasner, "A notation for infinite manifolds," *American Mathematical Monthly*, vol. 49 (1942), pp. 243-244.

Any transformation not of the union-preserving type converts at most $\infty^{(n^2+5n+2)/2}$ or ∞^∞ unions into unions. For $n=1$, this coincides with a theorem of Kasner concerning transformations of surface-elements of first order.

The only available union-preserving transformations in the whole domain of surface-elements are firstly, Lie's group of contact transformations of planar-elements, and secondly, our set of union-preserving transformations from surface-elements of order n into planar-elements, together with the extensions of these two types.

3. The conditions for union-preserving transformations. Any transformation T from surface-elements of order n into planar-elements is defined by equations of the forms

$$(1)\qquad \begin{aligned} &X=X(x,y,z;p_{j,m-j}), \quad Y=Y(x,y,z;p_{j,m-j}), \quad Z=Z(x,y,z;p_{j,m-j}),\\ &P=P(x,y,z;p_{j,m-j}), \quad Q=Q(x,y,z;p_{j,m-j}), \end{aligned}$$

where

$$(2)\qquad p_{j,m-j}=\partial^m z/\partial x^j\partial y^{m-j} \text{ for } j=0,1,2,\cdots,m, \text{ and } m=1,2,\cdots,n.$$

Of course, $P=\partial Z/\partial X$ and $Q=\partial Z/\partial Y$.

Thus any transformation T converts the totality of $\infty^{(n^2+3n+6)/2}$ surface-elements of order n of a certain region of the (x,y,z)-space into the ∞^5 planar-elements of a certain region of the (X,Y,Z)-space provided that the rank of the jacobian-matrix

$$(3)\qquad \begin{bmatrix} X_x & Y_x & Z_x & P_x & Q_x \\ X_y & Y_y & Z_y & P_y & Q_y \\ X_z & Y_z & Z_z & P_z & Q_z \\ X_{p_{j,m-j}} & Y_{p_{j,m-j}} & Z_{p_{j,m-j}} & P_{p_{j,m-j}} & Q_{p_{j,m-j}} \end{bmatrix}$$

which consists of five columns and $(n^2+3n+6)/2$ rows, is five throughout a certain region of the (x,y,z)-space.

A *double-series* consists of ∞^2 surface-elements of order n and may be defined by the parametric equations: $x=x(r,s)$, $y=y(r,s)$, $z=z(r,s)$; $p_{j,m-j}=p_{j,m-j}(r,s)$, where r and s are independent parameters, such that at least two of the functions are functionally independent. If all the functions in this system of equations are functionally related, the surface-elements of order n reduce to ∞^1 in number, and the resulting geometric configuration is called a *simple series.*

A double-series forms a *union* if and only if

$$(4)\qquad dz=p_{10}dx+p_{01}dy, \qquad dp_{j,m-j}=p_{j+1,m-j}dx+p_{j,m-j+1}dy,$$

for $j = 0, 1, 2, \cdots, m$, and $m = 1, 2, \cdots, n-1$. A simple series is termed a *strip* if the system of equations defining it satisfy the above system (4) identically.

A special type of union is the *conical-union*. This consists of ∞^2 surface-elements of order n which have in common a fixed surface-element of order $(n-1)$. The equations of any conical union are: $x = x_0$, $y = y_0$, $z = z_0$, $p_{10} = (p_{10})_0$, $p_{01} = (p_{01})_0, \cdots, p_{n-1,0} = (p_{n-1,0})_0, \cdots, p_{0,n-1} = (p_{0,n-1})_0$, $p_{n,0} = p_{n,0}(r, s), \cdots, p_{0,n} = p_{0,n}(r, s)$, where r and s are independent parameters. Through a given surface-element of order $(n-1)$, there pass $\infty^{(n-1)f(2)}$ conical-unions. Thus in totality, there are $\infty^{(n^2+n+4)/2+(n-1)f(2)}$ conical-unions.

Before continuing with our work, it is found convenient to introduce the two linear operators

$$
\begin{aligned}
A_k &= \partial/\partial x + p_{10}(\partial/\partial z) + \sum_{m=1}^{k} \sum_{j=0}^{m} p_{j+1,m-j}(\partial/\partial p_{j,m-j}), \\
B_k &= \partial/\partial y + p_{01}(\partial/\partial z) + \sum_{m=1}^{k} \sum_{j=0}^{m} p_{j,m-j+1}(\partial/\partial p_{j,m-j}).
\end{aligned} \tag{5}
$$

It follows that if λ is any function of $(x, y, z; p_{j,m-j})$ involving partial derivatives of order k at most, then $A_k(\lambda)$ (or $B_k(\lambda)$) represents the total partial derivative of λ with respect to x (or y) where z is considered to depend on both x and y. It is found that both $A_k(\lambda)$ and $B_k(\lambda)$ involve partial derivatives of order $(k+1)$, at most.

We seek all the unions that become unions under the transformation T as defined by the system of equations (1). By (1), (3), and (5), we find that the required unions satisfy the system of partial differential equations

$$
A_n(Z) = PA_n(X) + QA_n(Y), \qquad B_n(Z) = PB_n(X) + QB_n(Y), \tag{6}
$$

of order $(n+1)$. In general, there are at most $\infty^{(n^2+5n+2)/2}$ solutions except when the equations are in involution in which case the solutions involve arbitrary functions.[6]

THEOREM 1. *All the union-preserving transformations T from surface-elements of order n into planar-elements satisfy the $(n+3)$ conditions*

$$
\begin{aligned}
Z_{p_{j,n-j}} &= PX_{p_{j,n-j}} + QY_{p_{j,n-j}} \text{ for } j = 0, 1, 2, \cdots, n, \\
A_{n-1}(Z) &= PA_{n-1}(X) + QA_{n-1}(Y), \\
B_{n-1}(Z) &= PB_{n-1}(X) + QB_{n-1}(Y).
\end{aligned} \tag{7}
$$

---

[6] Kasner, "General transformation theory of differential elements," *American Journal of Mathematics*, vol. 32 (1904), pp. 392-401.

Any other transformation T converts either at most $\infty^{(n^2+5n+2)/2}$ or ∞^∞ unions into unions.

In a later section, we shall prove that all union-preserving transformations may be defined by a single directrix equation or a pair or triplet of such directrix equations.

4. The degenerate union-preserving transformations. A union-preserving transformation T is said to be nondegenerate if the rank of the jacobian-matrix

$$\begin{bmatrix} X_x & Y_x & Z_x \\ X_y & Y_y & Z_y \\ X_z & Y_z & Z_z \\ X_{p_{j,m-j}} & Y_{p_{j,m-j}} & Z_{p_{j,m-j}} \end{bmatrix}, \tag{8}$$

formed from the first three columns of the jacobian-matrix (3), is three in a certain region of the (x, y, z)-space. If the rank is two or less our union-preserving transformation T is called degenerate. Thus a union-preserving transformation T is called degenerate if and only if there exists at least one functional relationship between the components (X, Y, Z). We shall give a very brief discussion of the degenerate cases.

If all three functions (X, Y, Z) are identically constant, then T converts every union into a conical-union; these conical-unions all pass through a common point. If only two of the functions (X, Y, Z) are identically constant, then either $P = 0$ or $Q = 0$ or $1/P = 0$ or $1/Q = 0$, in which case every union is converted into a strip whose points describe a fixed line parallel to one of the coordinate axes. Finally if only one of the functions is identically constant, then either $P = Q = 0$, or $P/Q = 1/Q = 0$, or $Q/P = 1/P = 0$, in which case every union is converted into a plane parallel to one of the coordinate planes.

If there exist exactly two functional relationships between (X, Y, Z), then every union is carried into a single fixed strip.

Finally if there is a single functional relationship between the three components (X, Y, Z), then every union is carried into a single fixed surface.

Henceforth we shall consider only nondegenerate union-preserving transformations.

5. The general union-preserving transformations. We shall say that a non-degenerate union-preserving transformation T is general if it does not convert every conical-union of order n into a strip or conical-union. It is

said to be *intermediate* if every conical-union is carried into a strip. Finally we shall term T *special* if every conical-union is sent into a conical-union. Thus a non-degenerate transformation T of the union-preserving type is either general, or intermediate, or special.

A union-preserving transformation T is general, or intermediate, or special according as the rank of the jacobian-matrix

$$(9) \qquad (X_{p_{j,n-j}}, \quad Y_{p_{j,n-j}} \quad Z_{p_{j,n-j}}) \text{ for } j = 0, 1, 2, \cdots, n,$$

is of rank two, or one, or zero in a certain region of the (x, y, z)-space. Of course, by Theorem 1, the rank of this matrix cannot be three.

In this section, we shall study general union-preserving transformations so that the rank of the jacobian-matrix (9) is two.

THEOREM 2. *For a general union-preserving transformation T, the $(n+1)$ independent variables $p_{j,n-j}$ for $j = 0, 1, 2, \cdots, n$, can be eliminated from the three functions (X, Y, Z); thus obtaining the single eliminant*

$$(10) \qquad \Omega(X, Y, Z, x, y, z, p_{10}, p_{01}, \cdots, p_{n-1,0}, \cdots, p_{0,n-1}) = 0.$$

We call this the *directrix equation* of our general union-preserving transformation T.

The above result is an immediate consequence of the first $(n+1)$ equations (7). For clearly the jacobian-matrix (9) is of rank two. Thus by the theory of jacobians, we find that all the partial derivatives of order n may be eliminated from the three functions (X, Y, Z) yielding the single directrix equation (10).

THEOREM 3. *A general union-preserving transformation is determined completely by a single directrix equation.*

Thus a general union-preserving transformation T depends on one arbitrary function of $(n^2+n+8)/2$ independent variables. The content of the set of general union-preserving transformations is $\infty^{1f[(n^2+n+8)/2]}$.

In the first place, it is obvious that under a general union-preserving transformation T, any point in the (X, Y, Z)-space corresponds to a family of ∞^∞ surfaces in the (x, y, z)-space, defined by a partial differential equation of order $(n-1)$. These surfaces are defined by the directrix equation (10) where (X, Y, Z) are considered as constants. The partial derivatives of these surfaces of orders n at most, satisfy the equations

$$(11) \qquad \Omega = 0, \qquad A_{n-1}(\Omega) = 0, \qquad B_{n-1}(\Omega) = 0.$$

Secondly, any conical-union of order n in the (x, y, z)-space is converted

into a union in the (X, Y, Z)-space. This union is defined by the directrix equation (10) where $(x, y, z, p_{10}, p_{01}, \cdots, p_{n-1,0}, \cdots, p_{0,n-1})$ are regarded as constants. The partial derivatives of first order of this union satisfy the two equations

$$\Omega_X + P\Omega_Z = 0, \qquad \Omega_Y + Q\Omega_Z = 0. \tag{12}$$

We shall prove that the functions (X, Y, Z, P, Q) of our general union-preserving transformation T must satisfy the equations (11) and (12). It is seen that the components (X, Y, Z) of T satisfy the directrix equation (10) since it was obtained as a result of eliminating the partial derivatives of order n: $p_{j,n-j}$ from the first three of equations (1) defining T.

Next let us differentiate the directrix equation (10) with respect to $p_{j,n-j}$. Upon using the first $(n+1)$ conditions of (7), we find that the result is

$$X_{p_{j,n-j}}(\Omega_X + P\Omega_Z) + Y_{p_{j,n-j}}(\Omega_Y + Q\Omega_Z) = 0. \tag{13}$$

Since for a general union-preserving transformation T, the rank of the jacobian-matrix (9) is two, it follows from (13) that we deduce the equations (12).

Apply the linear operators (5) where $k = n - 1$ to the directrix equation (10) where (X, Y, Z) are given by the first three of equations (1). Because equations (12) are satisfied, we obtain the equations (11).

That we can actually solve (11) and (12) for (X, Y, Z, P, Q) follows from the fact that the directrix equation (10) represents a three-parameter family of partial differential equations of order $(n-1)$ in the (x, y, z)-space.

6. The intermediate union-preserving transformations. In this section, we shall consider the union-preserving transformations T whereby every conical-union of order n is converted into a strip. Thus the rank of the jacobian-matrix (9) is one.

THEOREM 4. *For an intermediate union-preserving transformation T, the $(n+1)$ independent variables $p_{j,n-j}$ for $j = 0, 1, 2, \cdots, n$, can be eliminated from the three functions (X, Y, Z), thus obtaining the two eliminants*

$$\begin{aligned} &\Omega_1(X, Y, Z, x, y, z, p_{10}, p_{01}, \cdots, p_{n-1,0}, \cdots, p_{0,n-1}) = 0, \\ &\Omega_2(X, Y, Z, x, y, z, p_{10}, p_{01}, \cdots, p_{n-1,0}, \cdots, p_{0,n-1}) = 0. \end{aligned} \tag{14}$$

These are the two directrix equations of any intermediate union-preserving transformation T.

The above result follows from the fact that the rank of the jacobian-matrix (9) is unity.

THEOREM 5. *The five functions* (X, Y, Z, P, Q) *defining an intermediate union-preserving transformation* T *satisfy not only the two directrix equations* (14) *but also the three equations*

$$(15) \qquad \frac{\Omega_{1X} + P\Omega_{1Z}}{\Omega_{2X} + P\Omega_{2Z}} = \frac{\Omega_{1Y} + Q\Omega_{1Z}}{\Omega_{2Y} + Q\Omega_{2Z}} = \frac{A_{n-1}(\Omega_1)}{A_{n-1}(\Omega_2)} = \frac{B_{n-1}(\Omega_1)}{B_{n-1}(\Omega_2)}.$$

The transformation T *is obtained by solving* (14) *and* (15) *for* (X, Y, Z, P, Q).

Thus a pair of directrix equations (14) determines completely an intermediate union-preserving transformation T. Any such transformation T is determined by two arbitrary functions of $(n^2 + n + 6)/2$ independent variables. The content of the set of intermediate union-preserving transformations is $\infty^{2f[(n^2+n+6)/2]}$.

To prove our Theorem 5, we proceed in the following manner. First of all, the components (X, Y, Z) satisfy the two directrix equations (14) since they were obtained as a result of eliminating the partial derivatives of order n: $p_{j,n-j}$ from the first three of equations (1) defining T. Moreover, these two directrix equations (14) are functionally independent.

Next differentiate each of the directrix equations (14) with respect to $p_{j,n-j}$. Introducing the first $(n+1)$ conditions (7), we obtain the results

$$(16) \qquad \begin{aligned} X_{p_{j,n-j}}(\Omega_{1X} + P\Omega_{1Z}) + Y_{p_{j,n-j}}(\Omega_{1Y} + Q\Omega_{1Z}) &= 0, \\ X_{p_{j,n-j}}(\Omega_{2X} + P\Omega_{2Z}) + Y_{p_{j,n-j}}(\Omega_{2Y} + Q\Omega_{2Z}) &= 0. \end{aligned}$$

Since the rank of the jacobian-matrix (9) is unity, the ratios $X_{p_{j,n-j}}/Y_{p_{j,n-j}}$ are all equal for $j = 0, 1, 2, \cdots, n$. Also not all of the expressions $X_{p_{j,n-j}}$ and $Y_{p_{j,n-j}}$ are identically zero. For otherwise by (7), the rank of the jacobian-matrix (9) would be zero. Hence from these remarks, we find that the equations (16) yield only a single eliminant which is equivalent to the first proportion of (15).

Apply the linear operators (5) where $k = n - 1$ to the directrix equations (14) where (X, Y, Z) are given by the first three of equations (1). Because of these conditions (7), we find

$$(17) \qquad \begin{aligned} (\Omega_{1X} + P\Omega_{1Z})A_{n-1}(X) + (\Omega_{1Y} + Q\Omega_{1Z})A_{n-1}(Y) + A_{n-1}(\Omega_1) &= 0, \\ (\Omega_{2X} + P\Omega_{2Z})A_{n-1}(X) + (\Omega_{2Y} + Q\Omega_{2Z})A_{n-1}(Y) + A_{n-1}(\Omega_2) &= 0, \end{aligned}$$

and a similar set of two equations where the operator A_{n-1} is replaced by the operator B_{n-1}.

Because of the first proportion of (15) which was established in the earlier paragraphs, it follows that for the equations (17) and the similar set where the operator A_{n-1} is replaced by the operator B_{n-1}, to be valid, we must have the continued proportion (15).

It is assumed that Ω_1 and Ω_2 are so constructed that five of the equations derived from (14) and (15) are functionally independent with respect to (X, Y, Z, P, Q). Hence these can be solved actually for (X, Y, Z, P, Q); thus obtaining our intermediate union-preserving transformation T.

7. The special union-preserving transformations. A union-preserving transformation is special if it converts every conical-union of order n into a conical-union of planar-elements. In this event, the components (X, Y, Z) will not involve the partial derivatives of order n: $p_{j,n-j}$, explicitly. Thus the first $(n+1)$ conditions of (7) are satisfied identically. The P and Q of such a transformation are determined by the remaining two equations of (7).

From the above remarks, it is seen that *a special union-preserving transformation is determined completely by a system of three functionally independent directrix equations of the forms* (10).

The special union-preserving transformations depend on three arbitrary functions of $(n^2+n+4)/2$ independent variables. The content of the set of special union-preserving transformations is $\infty^{3[f\,(n^2+n+4)/2]}$.

This completes our discussion of the union-preserving transformations. The non-degenerate union-preserving transformations are general, intermediate, or special. *These are determined by a single, pair, or triplet of directrix equations, respectively.* The contents of the three different classes are $\infty^{1f[(n^2+n+8)/2]}$, $\infty^{2f[(n^2+n+6)/2]}$, $\infty^{3f[(n^2+n+4)/2]}$, respectively.

It is noticed how the above result constitutes a direct extension of Lie's theorem on contact transformations.[7]

8. The union-preserving transformations in the domain of surface-elements. The following result concerning union-preserving transformations from surface-elements of order $n \geqq 2$ into surface-elements of second order, will be established.

---

[7] Kasner, "Lineal element transformations of space for which normal congruences of curves are converted into normal congruences," *Duke Mathematical Journal*, vol. 5 (1939), pp. 72-83. Also the following two papers by Kasner and De Cicco: "Curvature element transformations which preserve integrable fields," *Proceedings of the National Academy of Sciences*, vol. 25 (1939), pp. 104-111; and "Transformation theory of integrable double-series of lineal elements," *Bulletin of the American Mathematical Society*, vol. 46 (1940), pp. 93-100.

THEOREM 6. *All the non-degenerate union-preserving transformations from surface-elements of order $n \geqq 2$ into surface-elements of second order, are the extensions of the union-preserving transformations of surface-elements of order $(n-1)$ into planar-elements.*

A non-degenerate transformation from surface-elements of order $n \geqq 2$: $(x, y, z; p_{j,m-j})$ for $j = 0, 1, 2, \cdots, m$ and $m = 1, 2, \cdots, n$, into surface-elements of second order: (X, Y, Z, P, Q, R, S, T) is given by equations of the forms

$$\begin{aligned} &X = X(x, y, z; p_{j,m-j}), \quad Y = Y(x, y, z; p_{j,m-j}), \quad Z = Z(x, y, z; p_{j,m-j}), \\ &P = P(x, y, z; p_{j,m-j}), \quad Q = Q(x, y, z; p_{j,m-j}), \\ &R = R(x, y, z; p_{j,m-j}), \quad S = S(x, y, z; p_{j,m-j}), \quad T = T(x, y, z; p_{j,m-j}), \end{aligned} \tag{18}$$

where the jacobian-matrix of these eight functions is of rank eight in a certain region of the (x, y, z)-space.

We seek to find all those unions of surface-elements of order n which are converted into unions of surface-elements of second order by this transformation (18). The differential conditions (4) must correspond to the conditions

$$dZ = PdX + QdY, \quad dP = RdX + SdY, \quad dQ = SdX + TdY. \tag{19}$$

Thus by (5), we find the equations

$$\begin{aligned} &A_n(Z) = PA_n(X) + QA_n(Y), \quad B_n(Z) = PB_n(X) + QB_n(Y), \\ &A_n(P) = RA_n(X) + SA_n(Y), \quad B_n(P) = RB_n(X) + SB_n(Y), \\ &A_n(Q) = SA_n(X) + TA_n(Y), \quad B_n(Q) = SB_n(X) + TB_n(Y), \end{aligned} \tag{20}$$

of order $(n+1)$. In general, there are at most $\infty^{(n^2+5n-6)/2}$ or ∞^∞ unions which become unions.

From the above equations, we discover that if the transformation (18) is union-preserving, it must satisfy the conditions

$$\begin{aligned} Z_{p_{j,n-j}} &= PX_{p_{j,n-j}} + QY_{p_{j,n-j}}, \quad A_{n-1}(Z) = PA_{n-1}(X) + QA_{n-1}(Y), \\ &\qquad B_{n-1}(Z) = PB_{n-1}(X) + QB_{n-1}(Y), \\ P_{p_{j,n-j}} &= RX_{p_{j,n-j}} + SY_{p_{j,n-j}}, \quad A_{n-1}(P) = RA_{n-1}(X) + SA_{n-1}(Y), \\ &\qquad B_{n-1}(P) = RB_{n-1}(X) + SB_{n-1}(Y), \\ Q_{p_{j,n-j}} &= SX_{p_{j,n-j}} + TY_{p_{j,n-j}}, \quad A_{n-1}(Q) = SA_{n-1}(X) + TA_{n-1}(Y), \\ &\qquad B_{n-1}(Q) = SB_{n-1}(X) + TB_{n-1}(Y). \end{aligned} \tag{21}$$

Let us assume that not all the $X_{p_{j,n-j}}$ and $Y_{p_{j,n-j}}$ are zero. Accordingly our union-preserving transformation (18) is general or intermediate according

as the rank of the jacobian-matrix $(X_{p_{j,n-j}}, Y_{p_{j,n-j}})$ is two or one in a certain region of the (x, y, z)-space.

First consider the general case. From (X, Y, Z) of (18), we may eliminate all the partial derivatives of order n, obtaining a single eliminant of the form

$$Z = F(X, Y, x, y, z; p_{j,m-j}) \tag{22}$$

where $j = 0, 1, 2, \cdots, m$, and $m = 1, 2, \cdots, n-1$.

Substituting this into the conditions (21), we find the following results

$$\begin{gathered} P = F_X, \quad Q = F_Y, \quad R = F_{XX}, \quad S = F_{XY}, \quad T = F_{YY}, \\ A_{n-1}(F) = 0, \quad B_{n-1}(F) = 0, \quad A_{n-1}(F_X) = 0, \\ B_{n-1}(F_X) = 0, \quad A_{n-1}(F_Y) = 0, \quad B_{n-1}(F_Y) = 0. \end{gathered} \tag{23}$$

Thus the components (X, Y, Z, P, Q, R, S, T) of our transformation (18) must satisfy the equations (22) and (23).

Upon differentiating the conditions $A_{n-1}(F) = 0$, $B_{n-1}(F) = 0$, with respect to the $p_{j,n-j}$, we find, in view of the last four of equations (23), the conditions $F_{p_{j,n-1-j}} = 0$. Hence our transformation (18) must satisfy the conditions

$$A_{n-2}(F) = 0, \quad B_{n-2}(F) = 0, \quad F_{p_{j,n-1-j}} = 0. \tag{24}$$

At least one of the components (X, Y) must be effectively present in these equations. For otherwise these equations are identities and thus our eliminant (22) is of the form $Z = F(X, Y)$. This contradicts the fact that our transformation (18) is not degenerate.

Next it is seen that this system of equations (24) is not satisfied by a single solution of the form: $Y = Y(X, x, y, z; p_{j,m-j})$, (or $X = X(Y, x, y, z; p_{j,m-j})$), where $j = 0, 1, 2, \cdots m$, and $m = 1, 2, \cdots, n-1$. For then our transformation (18) is defined by two eliminants and this contradicts the fact that the transformation (18) is general.

Thus it follows that the system (24) must have a common solution of the form: $X = X(x, y, z; p_{j;m-j})$, $Y = Y(x, y, z; p_{j;m-j})$, which involves partial derivatives of order $n-1$, at most. This contradicts the fact that our transformation (18) is general.

Therefore we have succeeded in establishing that there are no general union-preserving transformations from surface-elements of order $n \geqq 2$ into surface-elements of second order.

Finally consider the intermediate case. From (X, Y, Z) of (18), we can eliminate the partial derivatives of order n, obtaining two functionally independent eliminants of the forms

$$(25) \qquad Y = F(X, x, y, z; p_{j,m-j}), \qquad Z = G(X, x, y, z; p_{j,m-j}),$$

for $j = 0, 1, 2, \cdots, m$, and $m = 1, 2, \cdots, n-1$. These are found on the assumption that at least one $X_{p_{j,n-j}}$ is not zero. Otherwise a $Y_{p_{j,n-j}}$ is not zero and the roles of X and Y are interchanged in these equations (25).

Substitute (25) into those conditions of (21) which involve Z. We find

$$(26) \qquad G_X = P + QF_X, \qquad A_{n-1}(G) = QA_{n-1}(F), \qquad B_{n-1}(G) = QB_{n-1}(F).$$

It is observed that the equations $A_{n-1}(F) = 0$, and $B_{n-1}(F) = 0$ (and similarly when F is replaced by G) can not hold simultaneously. For otherwise we can differentiate these with respect to the $p_{j,n-j}$, and thereby deduce that $F_{p_{j,n-1-j}} = 0$. Thus the function X must satisfy the conditions

$$(27) \qquad A_{n-2}(F) = 0, \qquad B_{n-2}(F) = 0, \qquad F_{p_{j,n-1-j}} = 0.$$

If X is not present in these equations, they must be identities. The first of equations (25) reduces to the form $Y = F(X)$. This contradicts the fact that our transformation (18) is not degenerate.

In the other possibility, the system (27) must possess a common solution X, which can involve partial derivatives of order $n-1$, at most. This contradicts our assumption that at least one $X_{p_{j,n-j}}$ is not zero.

The above arguments demonstrate that $A_{n-1}(F)$ and $B_{n-1}(F)$ (and similarly when F is replaced by G) are not simultaneously zero. From these facts and (26), we find that Q has the form

$$(28) \qquad Q = H(X, x, y, z; p_{j,n-j}) = \frac{A_{n-1}(G)}{A_{n-1}(F)} = \frac{B_{n-1}(G)}{B_{n-1}(F)}.$$

Substitute H for Q into the equations of (21) which involve the partial derivatives of Q with respect to the $p_{j,n-j}$. We derive

$$(29) \qquad (H_X - S - TF_X)X_{p_{j,n-j}} + H_{p_{j,n-j}} = 0.$$

These prove that the total partial derivatives of H with respect to the $p_{j,n-j}$ are proportional to the corresponding ones of X. Hence in the function H of (28), we must have $j = 0, 1, 2, \cdots, m$, and $m = 1, 2, \cdots, n-1$.

By using the conditions that the $Q_{p_{j,n-j}}$ are proportional to the $X_{p_{j,n-j}}$, we find by (28) that

$$(30) \qquad Q = \frac{A_{n-2}(G)}{A_{n-2}(F)} = \frac{B_{n-2}(G)}{B_{n-2}(F)} = \frac{G_{p_{j,n-1-j}}}{F_{p_{j,n-1-j}}}.$$

If this sytem of equations possesses a common solution for X, then X can involve partial derivatives of order $(n-1)$, at most. This contradicts our assumption that at least one $X_{p_{j,n-j}}$ is not zero.

The only other remaining case to be considered is when the equations (30) are all identities. All the partial derivatives of G with respect to the small letters are proportional to the corresponding ones of F in this case. Therefore, all the small letters may be eliminated from the two equations (25), yielding an eliminant of the form: $\lambda(X, Y, Z) = 0$. This contradicts the fact that our transformation (18) is not degenerate.

Thus it has been established that there are no intermediate union-preserving transformations from surface-elements of order $n \geqq 2$ into surface-elements of second order.

Therefore the only union-preserving transformations of the required kind are the extensions of the union-preserving transformations from surface-elements of order $(n-1)$ into planar-elements. Theorem 6 is proved completely.

THEOREM 7. *The only union-preserving transformations from surface-elements of order $n \geqq 1$ into surface-elements of order m where $n \geqq m \geqq 1$, are firstly, the contact group of planar-elements of Lie, and secondly, the union-preserving transformation from surface-elements of order n, where n is 2 or more, into planar-elements, together with the extensions of these two types.*

For, by hypothesis, any two unions which have contact of order $n \geqq 2$ must be carried into two unions of second-order contact, at least. By Theorem 6, our correspondence must be consequently an extension of a union-preserving transformation from surface-elements of order $(n-1)$ into planar-elements.

Therefore if T is a union-preserving transformation from surface-elements of order n into surface-elements of order m where $n \geqq m \geqq 1$, T is the extension of order m of a Lie contact transformation if $n = m$; and if $n > m$, T is the extension of order $(m-1)$ of a union-preserving transformation from surface-elements of order $(n - m + 1)$ into planar-elements.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

A NEW DEFINITION OF THE GODEAUX SEQUENCE OF QUADRICS.*

By CHENKUO PA.

A sequence of quadrics has been defined by L. Godeaux [1] at a generic point of an analytic non-developable surface in ordinary space. The first of them, ϕ, is the quadric of Lie and any two consecutive quadrics ϕ_{n-1}, ϕ_n in the sequence touch each other at four points, which are in turn the characteristic points of both quadrics. The second quadric ϕ_1 has been found independently by B. Su, who called it the associate quadric.[2] Godeaux [3] has shown the equivalence of this quadric and the quadric ϕ_1. In connection with these investigations the present author [4] has generalized this quadric from the point of view of Finikoff.[5] Following Godeaux, the definition of a Godeaux sequence consists in representing in S_5 the asymptotic tangents of a surface, and consequently has an indirect interpretation in S_3. The object of this paper is two-fold; first, to give a geometrical significance of this sequence by means of certain linear complexes associated with a point of a surface, and second, to generalize this new definition to a W congruence so as to obtain a generalized sequence of quadrics.[6]

1. For the subsequent discussion, it is convenient to utilize the normal coordinate system of Cartan [7] at a generic point M of a surface. Let $\{M M_1 M_2 M_3\}$ be a tetrahedron of reference of such a system; we have, on putting

* Received August 24, 1945.

[1] L. Godeaux, "Sur les lignes asymptotiques d'une surface et l'espace réglée," *Bulletin de l'Académie royale de Belgique* (1930), pp. 812-826; (1928), pp. 31-41, or "La théorie des surfaces et l'espace réglé" (1934).

[2] Buchin Su, "On the surfaces whose asymptotic curves belong to linear complexes II," *Tôhoku Mathematical Journal*, vol. 40 (1935), pp. 433-448.

[3] L. Godeaux, "Remarques sur les quadriques associées aux points d'une surface," *Journal of the Chinese Mathematical Society*, vol. 2 (1937), pp. 1-5.

[4] Chenkuo Pa, "A generalization of associate quadrics of a surface," *American Journal of Mathematics*, vol. 66 (1944), pp. 115-121.

[5] S. Finikoff, "Sur les quadriques de Lie et les congruences de M. Demoulin," *Recueil Mathématique de Moscou*, vol. 37 (1930), pp. 48-97.

[6] Cf. L. Godeaux, "Sur une propriété de l'enveloppe de certaines familles de quadriques," *Rendiconti dei Lincei*, Ser. 6, vol. 11 (1930), pp. 54-58.

[7] Cf. S. Finikoff, *Comptes Rendus*, vol. 197 (1933), pp. 883-885; B. Su, *Tôhoku Mathematical Journal*, vol. 41 (1935), pp. 203-215.

$$| M M_1 | = U, \qquad | M M_2 | = V,$$

that U (or V) is the image of the asymptotic u-tangent (v-tangent) in the Klein hyperquadric Ω of S_5. The locus of U (or V) forms a conjugate net on Ω and U satisfies the following Laplace equation:

$$\frac{\partial^2 U}{\partial u \partial v} - \frac{\partial \log \beta}{\partial v} \frac{\partial U}{\partial v} - \beta\gamma U = 0.$$

Thus it gives rise to a Laplace sequence self-conjugate to Ω, namely,

$$\cdots (U_{-n}), \cdots, (U_{-1}), (U), (U_1), \cdots, (U_n), \cdots,$$

where $U_{-1} = V$ and (U_n) denotes the Laplace transform of (U_{n-1}) in the sense v. Let h_i, k_i be the Darboux-Laplace invariants of the net (U_i); then we have the following relations:

$$\frac{\partial U_n}{\partial v} = U_{n+1} + U_n \frac{\partial}{\partial v} \log (\beta h_1 \cdots h_n),$$

$$\frac{\partial U_n}{\partial u} = h_n U_{n-1}, \qquad (U_0 = U, \quad U_{-1} = V, \quad h_0 \equiv \beta),$$

$$\frac{\partial^2 U_n}{\partial u \partial v} - \frac{\partial U_n}{\partial v} \frac{\partial}{\partial v} \log (\beta h_1 \cdots h_n) - h_n U_n = 0;$$

$$\frac{\partial V_n}{\partial u} = V_{n+1} + V_n \frac{\partial}{\partial u} \log (\gamma k_1 \cdots k_n),$$

$$\frac{\partial V_n}{\partial v} = k_n V_{n-1}, \qquad (V_0 = V, \quad V_{-1} = U, \quad k_0 \equiv \gamma),$$

$$\frac{\partial^2 V_n}{\partial u \partial v} - \frac{\partial V_n}{\partial u} \frac{\partial}{\partial u} \log (\gamma k_1 \cdots k_n) - k_n V_n = 0.$$

The intersection of the plane $(U_n U_{n+1} U_{n+2})$ with Ω is known to be the image in S_5 of one regulus of the quadric ϕ_n, namely, the n-th quadric in the Godeaux sequence.

Let us now consider the osculating linear complex $R_1(u, v)$ of the asymptotic u-curve at the point $M(u, v)$. It can easily be shown that this complex is represented in S_5 by the intersection of the hyperplane $[U, V, V_1, V_2, V_3]$ with Ω. If $M(u + du, v + dv)$ be any point of S near $M(u, v)$, the corresponding linear complex $R_1(u + du, v + dv)$ may then be expressed by $R_1 + R_{1u}du + R_{1v}dv + \cdots$ and therefore all the linear complexes corresponding to points in the first order neighborhood of M have a common regulus formed by lines common to $R_1(u, v)$, $R_{1u}(u, v)$, $R_{1v}(u, v)$. *This regulus constitutes the quadric ϕ of Lie,* as was shown in the former paper.

We now consider the linear complexes $R_1(u, v+dv) = R_1 + R_{1v}dv + (1/2!)R_{1vv}dv^2 + \cdots$ corresponding to points in the second order neighborhood of M along an asymptotic v-curve. We find easily that *these linear complexes have one regulus of the quadric ϕ_1 in common, and the linear complexes* $R_1(u, v+dv) = R_1 + R_{1v}dv + (1/2!)R_{1vv}dv^2 + (1/3!)R_{1vvv}dv^3 + \cdots$ *corresponding to points in the third order neighborhood of M along an asymptotic v-curve have two common straight lines g_1, g_2,* whose images on Ω are the points of intersection of (V_2V_3) with Ω. When these lines g_1, g_2 vary in the direction u, we obtain two ruled surfaces G_1, G_2. *There exists a linear complex $R_2(u, v)$ having a contact of the third order with both G_1 and G_2 along the generators g_1, g_2 respectively.* In fact this complex $R_2(u, v) \equiv R_2$ is determined by $[V_2V_3V_4V_5V_6]$. With the aid of $R_2(u, v)$, we obtain, on making use of a similar method, the quadrics ϕ_2, ϕ_3, ϕ_4. More precisely, they are determined as common reguli of the three sets of linear complexes $R_2 + R_{2u}du + (1/2!)R_{2uu}du^2$, $R_2 + R_{2u}du + R_{2v}dv$ and $R_2 + R_{2v}dv + (1/2!)R_{2vv}dv^2$ respectively. If we consider all the linear complexes $R_2 + R_{2v}dv + (1/2!)R_{2vv}dv^2 + (1/3!)R_{2vvv}dv^3$ corresponding to points in the third order neighborhood along an asymptotic v-curve, we get two straight lines similar to g_1, g_2. Applying the process successively we can define the whole sequence.

Thus we have proved the following

THEOREM. *The Godeaux sequence of quadrics associated with a point of a surface can be simply defined with the aid of asymptotic osculating linear complexes.*

2. In the next place let us consider the asymptotic osculating ruled surface R_u generated by v-tangents along an asymptotic u-curve. The image of the osculating linear complex $S_1(u, v)$ of R_u at M is evidently the hyperplane $[VV_1V_2V_3V_4]$. *All the linear complexes $S_1 + S_{1u}du + (1/2!)S_{1uu}du^2$ corresponding to points of the surface in the second order neighborhood of M along an asymptotic u-curve have one regulus of the quadric ϕ in common; all the linear complexes $S_1 + S_{1u}du + S_{1v}dv$ corresponding to points of the surface in the first order neighborhood of M have one regulus of the quadric ϕ_1 in common, and all the linear complexes $S_1 + S_{1v}dv + (1/2!)S_{1vv}dv^2$ corresponding to points of the surface in the second order neighborhood of M along an asymptotic v-curve have their common lines in one regulus of the quadric ϕ_2.* Similarly, among the linear complexes $S_1 + S_{1v}du + (1/2!)S_{1vv}dv^2 + (1/3!)S_{1vvv}dv^3$ corresponding to points of the surface in the third order neighborhood of M along an asymptotic v-curve, there are only four inde-

pendent ones, namely, $S_1, S_{1v}, S_{1vv}, S_{1vvv}$ and, in consequence, *they have two common straight lines.* Proceeding in the same manner as in **1**, we can define the quadrics ϕ_3, ϕ_4, ϕ_5 successively, and thus obtain the whole sequence. In summary, we can state the following

THEOREM. *With the aid of osculating linear complexes of the asymptotic ruled surfaces associated with a point of a surface, we can also define the Godeaux sequence of quadrics.*

3. There is now no difficulty in generalizing the notion of a Godeaux sequence of quadrics to a W congruence.[8] For, by a classical theorem of Darboux, the image (U) of a W congruence in the Klein quadric of S_5 forms a conjugate net. Letting (U_{i+1}) be the Laplace transform of (U_i) in the sense v, we get a Laplace sequence

$$\cdots, (U_{-n-1}), (U_{-n}), \cdots, (U_{-1}), (U), (U_1), \cdots, (U_n), (U_{n+1}), \cdots.$$

Hence by the method used by Godeaux in the case of the asymptotic congruence, we can define a sequence of quadrics:

$$\cdots, \phi_{-n}, \cdots, \phi_{-2}, \phi_{-1}, \phi, \phi_1, \phi_2, \cdots, \phi_n, \cdots,$$

where ϕ_n is defined by $(U_nU_{n+1}U_{n+2})$.

It should be mentioned that the quadric ϕ_{-1} decomposes into two focal planes of the congruence, and this sequence has properties similar to the Godeaux sequence.

By the method used in **1** and **2**, we can readily define the new sequence. In the present case, the quadrics of the sequence can, however, be determined by the *osculating complexes* of the congruence.

In fact, the osculating linear complex of a W conguence is determined by $(U_{-2}U_{-1}UU_1U_2)$. By the method of **2**, we first define the quadrics ϕ_{-2}, ϕ_{-1}, ϕ, and then obtain, on using the same complex in the sense v, the quadrics ϕ_1, ϕ_2, ϕ_3, and similarly, the quadrics ϕ_{-3}, ϕ_{-4}, ϕ_{-5}, etc.

Thus we have arrived at the following

THEOREM. *With the aid of osculating complexes of a W congruence, a sequence of quadrics can be defined to be associated with each ray of the congruence, that is, the generalization of a Godeaux sequence of a surface to a W congruence.*

NATIONAL UNIVERSITY OF CHEKIANG,
KWEICHOW, MEITAN, CHINA.

[8] L. Godeaux, *loc. cit.* [6].

UNILATERAL VARIATIONS WITH VARIABLE END-POINTS.*

By JULIAN D. MANCILL.

1. Introduction. For the derivation of certain necessary conditions for an arc E_{12} to minimize an integral in the form

$$I = \int_{t_1}^{t_2} F(x, y, x', y')\, dt \equiv \int_{t_1}^{t_2} F(x, x')\, dt,$$

we shall assume that $F(x, x')$, $F_{x'}(x, x')$, and $F_{y'}(x, x')$ are continuous in a region R of the xy-plane, and for all $(x', y') \neq (0, 0)$. The function $F(x, x')$ is assumed to be positively homogenous of the first degree in x' and y'. Admissible curves will be supposed always in the form

$$x = x(t), \qquad y = y(t),$$

and of class D' in the region R.[1] In order to emphasis the problem at hand, *we shall assume, unless otherwise stated, that the minimizing curve E_{12} lies along the boundary of the region R of admissible curves.* Our problem then is to derive the properties of the arc E_{12} in order that it minimize the integral I in the class of all admissible curves joining a given curve D through the point 1 and a given curve C through the point 2. We shall assume that the curves D and C are of class C' in the region R.

The usual procedure in deriving necessary conditions in order that an arc minimize the integral I in the case of free variations has been to derive the Euler equations first [2] and then to make use of them in deriving further necessary conditions. Recent studies have shown that such a procedure is not necessary nor desirable for the derivation of two of the so called first order conditions. These two conditions for the arc E_{12} are the Weierstrass condition,[3] which states that

$$\mathcal{E}(x, x', p) \equiv F(x, p) - pF_{x'}(x, x') - qF_{y'}(x, x') \geqq 0$$

along E_{12} for all $(p, q) \neq (0, 0)$, and the corner conditions.[4]

---

* Received October 22, 1945; Revised September 5, 1946.

[1] For a definition of the term "class" as here used see Bolza, *Lectures on the calculus of variations*, University of Chicago, p. 7.

[2] See, e. g., Bolza, *loc. cit.*, p. 122.

[3] Graves, "A proof of the Weierstrass condition in the calculus of variations," *American Mathematical Monthly*, vol. 41 (1934), pp. 502-504.

[4] Mancill, "A proof of the corner conditions in the calculus of variations," *American Mathematical Monthly*, vol. 43 (1936), pp. 68-70.

In the present paper we shall derive a third first order condition, the transversality condition at the points 1 and 2, without making use of the Euler equations. These transversality conditions involve inequalities in the case of unilateral variations instead of equalities as in the case of free variations. The proof is made with weaker hypotheses than the usual proof by means of the first variation. The method of proof is simple and may be applied to parametric or non-parametric problems in space of any number of dimensions. So far as the author knows these results are new since there seems to be no reference in the literature to transversality at points of the boundary of the region of admissible curves.

Jacobi's criterion will be formulated for the cases in which the Euler equations are satisfied along E_{12}. There is no restriction on the length of E_{12} in the case where the Euler equations are not satisfied along E_{12}.

A set of sufficient conditions in order that the arc E_{12} shall minimize the integral I will be given after appropriate modification of the original hypotheses on the region R and the function $F(x, x')$. There are many cases and they lead to a large number of distinct sets of sufficient conditions. The only cases not treated fully are the ones usually omitted, that is when $F(x, x') = 0$ along E_{12} at the points 1 and 2 when the curves D and C are transversal or tangent to E_{12}. The methods used are applicable to non-parametric problems.

The treatment of sufficient conditions given here for the case where E_{12} is an extremal and the curves D and C are transversal to E_{12} applies to free variations, and the method of proof is simpler than that given by Bliss [5] for non-parametric problems. He assumes that the integrand function is positive throughout the region of admissible curves for all directions, whereas in the present treatment it is only assumed that $F(x, x') \neq 0$ along E_{12} at the points 1 and 2.

2. The tranversality condition. Let the equations of the minimizing curve E_{12} be

$$x = x(t), \qquad y = y(t), \qquad t_1 \leq t \leq t_2,$$

and those of the curves D and C be respectively

$$D: \quad x = X_1(\tau), \qquad y = Y_1(\tau),$$

and

$$C: \quad x = X_2(\tau), \qquad y = Y_2(\tau),$$

[5] "Jacobi's criterion when both end-points are variable," *Mathematische Annalen*, vol. 58 (1904), pp. 70-80.

interior to the region R. Let the equations

$$\mathcal{E}: \quad x = \phi(s, a), \qquad y = \psi(s, a)$$

represent a one parameter family of admissible curves intersecting E_{12} for $s = s_3(a)$, $t = t(a)$, and intersecting the curve C for $s = s_4(a)$, $\tau = \tau(a)$ in a neighborhood of the point 2. The incidence relations of these curves are expressed by the equations

$$(1) \qquad \begin{aligned} x[t(a)] &= \phi[s_3(a), a], \qquad y[t(a)] = \psi[s_3(a), a], \\ X_2[\tau(a)] &= \phi[s_4(a), a] \qquad Y_2[\tau(a)] = \psi[s_4(a), a], \end{aligned}$$

where we may assume $t'(a) > 0$ at $a = a_2$.

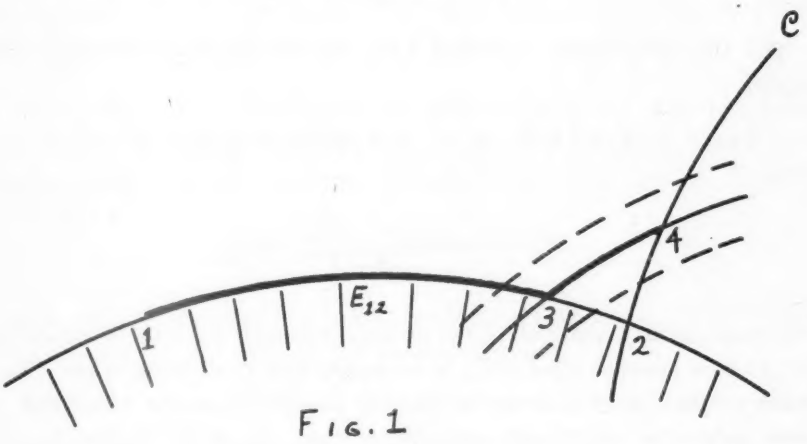

FIG. 1

Now consider the admissible comparison curve

$$E_a \equiv E_{13} + \mathcal{E}_{34},$$

of the type indicated by the heavy line in the figure. The function

$$I(a) = I(E_a)$$

has the derivative

$$(2) \qquad I'(a_2) = F(x, x')t'(a_2) - F(x, \phi')s'_3(a_2) + F(x, \phi')s'_4(a_2)$$

at $a = a_2$. By means of the homogeneity property of F, the incidence relations (1), and their derivatives with respect to a, we find that

$$(3) \qquad I'(a_2) = \mathcal{E}(x, \phi', x')t'(a_2) + [X'_2F_{x'}(x, \phi') + Y'_2F_{y'}(x, \phi')]\tau'(a_2) \leqq 0$$

at the point 2 since E_{12} is to minimize the integral I and $t'(a_2)$ is assumed positive. The parameter t may be chosen so that $t'(a_2) = 1$. Now, if we let the direction (ϕ', ψ') approach the direction (x', y') at 2 and assume that C is not tangent to E_{12} at 2, it follows from the inequality (3) that

$$(4)_2 \qquad X'_2F_{x'}(x, x') + Y'_2F_{y'}(x, x') \geqq 0$$

if $\tau'(a_2) < 0$, that is if the curve C is interior to R for $\tau \geqq \tau_2$. Likewise if $\tau'(a_2) > 0$, that is if C is interior to R for $\tau \leqq \tau_2$, it follows from (3) that

$$(5)_2 \qquad X'_2F_{x'}(x, x') + Y'_2F_{y'}(x, x') \leqq 0.$$

The two inequalities $(4)_2$ and $(5)_2$ can be combined into the single condition

$$\tilde{T}(x, x', X'_2) \equiv [X'_2F_{x'}(x, x') + Y'_2F_{y'}(x, x')]/D(x', X'_2) \geqq 0,$$

if

$$D(x', X'_2) \equiv \begin{vmatrix} x' & X'_2 \\ y' & Y'_2 \end{vmatrix} \neq 0,$$

as the transversality condition at the variable end-point 2. This is true because $D(x', X'_2)$ is positive when $(4)_2$ is necessary and is negative when $(5)_2$ is necessary. This condition can be derived directly from the inequality (3) if we make the additional assumption that the well known function $F_1(x, x') > 0$ along E_{12} at the point 2. For then it follows that

$$[X'_2F_{x'}(x, x') + Y'_2F_{y'}(x, x')]\tau'(a_2)/t'(a_2) < 0$$

at the point 2 for the direction (ϕ', ψ') sufficiently near but distinct from the direction (x', y') of E_{12} at the point 2, since [6] $\mathcal{E}(x, \phi', x') > 0$ at 2 under these conditions. From the derivatives of the relations (1), we have

$$\tau'(a_2)/t'(a_2) = D(\phi', x')/D(\phi', X'_2).$$

Since the admissible region R is to the left along E_{12}, the determinant $D(\phi', x')$ is negative and we have

[6] *Cf.* Bolza, *loc. cit.*, p. 140.

$$[X'_2F_{x'}(x,\phi') + Y'_2F_{y'}(x,\phi')]/D(\phi', X'_2) > 0$$

at the point 2 if the direction (ϕ', ψ') is sufficiently near but distinct from the direction (x', y') there. Therefore, it follows that $\tilde{T}(x, x', X'_2) \geqq 0$ if $D(x', X'_2) \neq 0$ at 2.

If $D(x', X'_2) = 0$ at the point 2, that is if $(X'_2, Y'_2) = (kx', ky')$ for $k \neq 0$, several cases arise:

1) If the curve C precedes the normal line to E_{12} at 2, then $s'_3(a_2) = s'_4(a_2)$ in (2) as follows from the derivatives of the relations (1). Therefore, in this case we have

$$F(x, x') \leqq 0,$$

since we assumed that $t'(a_2) > 0$. This result holds for k positive or negative and is independent of the element (x, ϕ') used in the construction.

2) If the curve C lies beyond the normal line to E_{12} at 2, then we may make use of the family of admissible curves $(\mathcal{E})$ following 2 and in the relation analogous to (2) we would have again $s'_3(a_2) = s'_4(a_2)$. Therefore, in this case

$$I'(a_2)/t'(a_2) = F(x, x') \geqq 0$$

at the point 2. This result is true for k positive or negative and is independent of the element (x, ϕ') used in the construction.

3) If the curve C interior to R lies on both sides of the normal line to E_{12} at 2, that is if C does not cross E_{12}, then the results of cases 1) and 2) must hold simultaneously and therefore in this case

$$F(x, x') = 0.$$

Analogous conditions can be shown to hold at the point 1 between the curves E_{12} and D *with each inequality reversed.*

If either of the points 1 or 2 is interior to the region R of admissible curves, then a construction similar to the one used above but on the opposite side of E_{12} shows, with the results already obtained, that for free variations all inequalities in the conditions are to be replaced by equalities. That is, at an interior variable end-point

$$\tilde{T}(x, x', X') = 0,$$

if the transversal direction (X', Y') is distinct from the direction (x', y'), and

$$F(x, x') = 0$$

if $(X', Y') = (kx', ky')$ for $k \neq 0$.

We may summarize the results of this section in the following:

THEOREM 1. *If the curve E_{12} minimizes the integral I in the class of all admissible curves joining the curves D and C; and*

1) *if the point 2 is interior to R, then $T(x, x', X'_2) = 0$ for $(X'_2, Y'_2) \neq (kx', ky')$ $(k > 0)$ and $F(x, x') = 0$ at 2 for $(X'_2, Y'_2) = (kx', ky')$ $(k \neq 0)$;*

2) *if the arc E_{12} is along the boundary of R, then $T(x, x', X'_2) \geqq 0$ at 2 for $(X'_2, Y'_2) \neq (kx', ky')$ $(k > 0)$ and for $(X'_2, Y'_2) = (kx', ky')$ $(k \neq 0)$*

$F(x, x') \leqq 0$ if C precedes the normal line to E_{12} at 2,

$F(x, x') \geqq 0$ if C lies beyond the normal to E_{12} at 2,

$F(x, x') = 0$ if C lies on both sides of the normal to E_{12} at 2.

Similar results hold at the point 1 for E_{12} and the curve D but with the inequalities reversed.

The conditions in the second part of Theorem 1 have a simple geometric interpretation. Suppose that the curve C is interior to R for $\tau \geqq \tau_2$. Then the inequality $(4)_2$ must hold even when $(X'_2, Y'_2) = (kx', ky')$ $(k \neq 0)$ since in this case

$$X'_2 F_{x'}(x, x') + Y'_2 F_{y'}(x, x') = kF(x, x') \geqq 0$$

at 2. Similarly, if the curve D is interior to R for $\tau \geqq \tau_1$, we have

$$(4)_1 \qquad X'_1 F_{x'}(x, x') + Y'_1 F_{y'}(x, x') \leqq 0$$

at the point 1. Consider now the two relations $(4)_1$ and $(4)_2$ along with the equations

$$(6) \qquad x' F_{x'}(x, x') + y' F_{y'}(x, x') = F(x, x'),$$

and

$$(7) \qquad X' F_{x'}(x, x') + Y' F_{y'}(x, x') = 0,$$

where (X', Y') denotes the direction transversal to (x', y'). The equations (6) and (7) in general involve three directions at a point of E_{12}, namely

(x', y'), (X', Y'), and $(F_{x'}, F_{y'})$, where we assume for the moment that not both the derivatives $F_{x'}$ and $F_{y'}$ are zero. Equation (7) shows that the direction (X', Y') is orthogonal to $(F_{x'}, F_{y'})$. Equation (6) shows that the angle between the directions (x', y') and $(F_{x'}, F_{y'})$ is acute if $F(x, x')$ is positive and is obtuse if $F(x, x')$ is negative. The inequality $(4)_2$ shows that the angle at 2 formed by the directions (X'_2, Y'_2) and $(F_{x'}, F_{y'})$ is acute, while

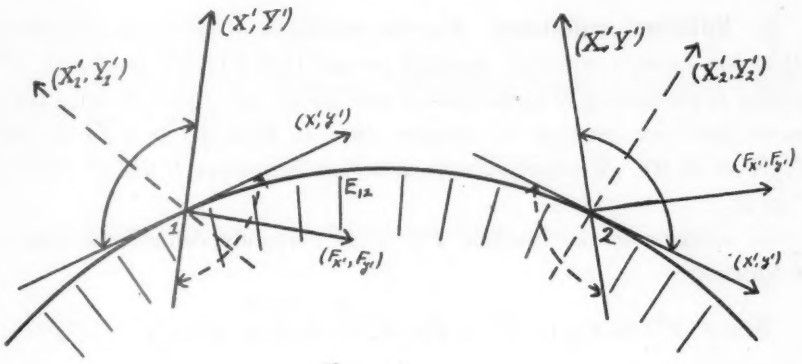

FIG. 2

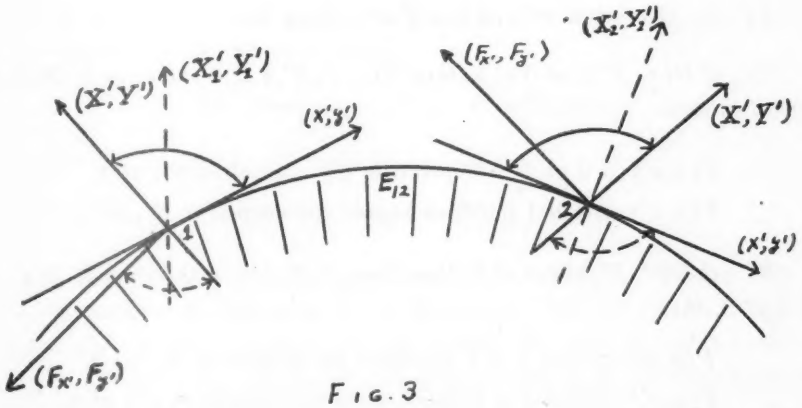

FIG. 3

the relations $(4)_1$ shows that the angle at 1 formed by the directions (X'_1, Y'_1) and $(F_{x'}, F_{y'})$ is obtuse. These relations are shown in figure 2 when F is positive at 1 and 2 and in figure 3 when F is negative at the points 1 and 2. The curved arrow indicates the possible range of the directions of the curves D and C directed toward the interior of R, and analogously the broken curved arrow indicates the range of the directions of D and C directed toward the exterior of R in each case. It follows, therefore that *if* $F(x, x') \neq 0$ *at* 1 *and* 2, *then there exist directions for which the conditions of the Theorem are satisfied.*

If $F(x, x') = 0$ at 1 and 2, then equations (6) and (7) show that the directions (x', y') and (X', Y') are coincident or in opposite directions at 1 or 2, unless $F_{x'}$ and $F_{y'}$ are both zero. In the latter case any direction may be defined as transversal to (x', y'). If $F_{x'}$ and $F_{y'}$ are not both zero but $F(x, x')$ is zero, there may or may not exist directions distinct from (x', y') for which the conditions of the Theorem are satisfied.

3. Sufficient conditions. For the treatment of sufficient conditions, as well as for Jacobi's criterion, we shall assume that $F(x, x')$ is of class C'''' in a region R' containing R in its interior and for all $(x', y') \neq (0, 0)$. *We shall assume that our problem is regular, that is* $F_1(x, x') > 0$ *in* R *for all* $(x', y') \neq (0, 0)$. We shall suppose also that the curves D and C are of class C'' in R.

In addition to the function $\bar{T}(x, x', X')$ already defined, we shall need the function [7]

$$T(x, x', x'') \equiv F_{xy'}(x, x') - F_{yx'}(x, x') + F_1(x, x')(x'y'' - x''y').$$

THEOREM 2. *Suppose that* E_{12} *has the properties:*

1) E_{12} *is of class* C'' *and has* $T < 0$ *along it;*

2) *if* $D(x', X'_1) \neq 0$ *at* 1, *then* $\bar{T}(x, x', X'_1) \leqq 0$ *there; or if* $D(x', X'_1) = 0$ *at* 1, *then*

$F(x, x') > 0$ *at* 1 *if* D *precedes the normal to* E_{12} *at* 1,
$F(x, x') < 0$ *at* 1 *if* D *lies beyond the normal to* E_{12} *at* 1;

3) *if* $D(x', X'_2) \neq 0$ *at* 2, *then* $\bar{T}(x, x', X'_2) \geqq 0$ *there; or if* $D(x', X'_2) = 0$ *at* 2, *then*

$F(x, x') < 0$ *at* 2 *if* C *precedes the normal to* E_{12} *at* 2,
$F(x, x') > 0$ *at* 2 *if* C *lies beyond the normal to* E_{12} *at* 2.

Then the arc E_{12} *minimizes the integral* I *in the class of all admissible curves sufficiently near* E_{12} *joining the curves* D *and* C.

Since $T < 0$ along E_{12} it remains less than zero along the boundary of R in neighborhoods preceding the point 1 and following the point 2. We shall

[7] If E_{12} is of class C'' and minimizes the integral I, then $T \leqq 0$ along it. For a proof see, Bolza, *loc. cit.*, p. 148.

denote this extended arc by E_{56}, where the points 5 and 6 are sufficiently near 1 and 2 respectively. Then the one parameter family of extremals

$$\mathcal{E}: \quad x = \phi(t, a), \qquad y = \psi(t, a),$$

tangent to E_{56} for $t = t(a)$, $a_5 \leqq a \leqq a_6$, forms a field $\mathcal{F}_2$ adjoining E_{56} in R for

$$t(a) \leqq t \leqq t(a) + e_2, \qquad a_5 \leqq t \leqq a_6,$$

if e_2 is sufficiently small.[8] Also, the family of extremals $(\mathcal{E})$ forms a field $\mathcal{F}_1$ adjoining E_{56} in R for

$$t(a) - e_1 \leq t \leq t(a), \qquad a_5 \leq a \leq a_6,$$

if e_1 is sufficiently small. The functional determinant

$$\Delta(t, a) \equiv \begin{vmatrix} \phi' & \phi_a \\ \psi' & \psi_a \end{vmatrix}$$

is positive interior to $\mathcal{F}_1$, negative interior to $\mathcal{F}_2$, and is zero on E_{56}. Now choose a point 8 on E_{12} at which $F(x, x') \neq 0$. Let L represent a curve of class C' through the point 8 in the unique direction transversal to E_{12} there. The one parameter family of extremals

$$\bar{\mathcal{E}}: \quad x = \bar{\phi}(t, a), \qquad y = \bar{\psi}(t, a),$$

cut by L transversally forms a field $\mathcal{F}_3$ in R for

$$t_8 - e_3 \leqq t \leqq t_8 + e_3, \qquad a_8 \leqq a \leqq a_8 + d,$$

where e_3 and d are sufficiently small positive numbers. This family of extremals contains the extremal $\mathcal{E}_8$ of the family $(\mathcal{E})$ for $a = a_8$. Let the region in the xy-plane common to the regions $\mathcal{F}_1$ and $\mathcal{F}_2$ be denoted by $\mathcal{F}$. Let $V_{1'2'}$ denote any admissible comparison curve in $\mathcal{F}$ joining the curves D and C in the points $1'$ and $2'$, respectively, and distinct from E_{12}. Let 7 represent the *last* point along $V_{1'2'}$, in which $V_{1'2'}$ crosses $\mathcal{E}_8$ preceding 8. Also, let 9 denote the *first* point beyond 7 along $V_{1'2'}$ in which $V_{1'2'}$ crosses $\mathcal{E}_8$ beyond 8. Then it is easily seen that $\mathcal{F}$ can be so restricted that V_{79} lies in the region $\mathcal{F}_3$.

Now, by considering only the field $\mathcal{F}_1$ and making use of the assumption of regularity, it can be shown that[9]

---

[8] Bolza, *Vorlesungen über Variationsrechnung*, Berlin, 1909, pp. 403-05.

[9] Mancill, "The minimum of a definite integral with respect to unilateral varia-

$$I(V_{1'7} + \mathcal{E}_{78}) > I(\mathcal{E}_{1'3} + E_{38}). \tag{8}$$

By considering the field $\mathcal{F}_2$ it can be shown in the same way that

$$I(\mathcal{E}_{89} + V_{92'}) > I(E_{84} + \mathcal{E}_{42'}). \tag{9}$$

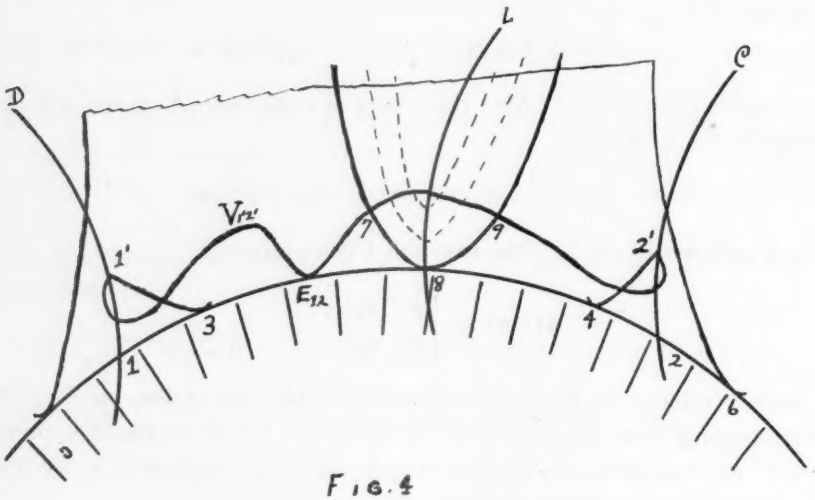

FIG. 4

Also, by making use of the well-known invariant integral I^* in the field $\mathcal{F}_3$, it can be shown that

$$I(V_{79}) > I(\mathcal{E}_{78} + \mathcal{E}_{89}). \tag{10}$$

Combining the inequalities (8), (9), and (10), we obtain

$$I(V_{1'2'}) > I(\mathcal{E}_{1'3} + E_{34} + \mathcal{E}_{42'}). \tag{11}$$

It is our purpose now to show that in general, under the hypotheses of our theorem,

$$I(\mathcal{E}_{1'3}) > I(E_{13}), \tag{12}$$

and

$$I(\mathcal{E}_{42'}) > I(E_{42}), \tag{13}$$

if the field $\mathcal{F}$ is sufficiently restricted.

---

tions," *Contributions to the Calculus of Variations*, 1933-1937, University of Chicago Press, pp. 129-32.

We shall prove the inequality (13) in two parts by making use of the extremals of the field $\mathfrak{F}_2$. First consider the case where $D(x', X'_2) \neq 0$ at the point 2. The function

$$I_2(a) \equiv I(E_{44'} + \mathcal{E}_{4'2''}),$$

where $4'$ is on E_{42} and $2''$ is on C, has the derivative

$$I'_2(a) = \Delta(t, a)\bar{T}(x, \phi', X'_2),$$

which is negative near the point 2 and is zero at 2, if $\bar{T}(x, x', X'_2) > 0$ there. If $\bar{T}(x, x', X'_2) = 0$ at 2 we shall need to consider the derivative of the function

$$g_2(a) \equiv [X'_2 F_{x'}(x, \phi') + Y'_2 F_{y'}(x, \phi')]^{2''}$$

near 2. Direct calculation shows that this function has the derivative

$$g'_2(a) = [A_2\Delta(t, a) + B_2\Delta_t(t, a)]/D(\phi', X'_2)$$

near 2, where A_2 is a known function of the elements of our problem and $B_2 = [D(\phi', X'_2)]^2 F_1(x, \phi')$. Therefore, $g'_2(a)$ has the same sign near 2 as has the function

$$D(x', X'_2)F_1(x, x')\Delta_t(t_2, a_2)$$

at 2; this function is not zero and has the sign opposite to that of $D(x', X'_2)$ since $\Delta_t(t_2, a_2)$ is negative. Therefore, it follows that even if $\bar{T}(x, \phi', X'_2) = 0$ at 2 it is positive near 2 in the field $\mathfrak{F}_2$, if this field is sufficiently restricted. Thus, $I'_2(a)$ is negative near 2 and the function $I_2(a)$ is decreasing. But it is obvious that

$$I(\mathcal{E}_{42'}) - I(E_{42}) = I_2(a_4) - I_2(a_2),$$

where $a_4 < a_2$ if $2'$ is distinct from 2, and the inequality (13) is proved.

Next we shall consider the case where $D(x', X'_2) = 0$, that is $(X'_2, Y'_2) = (kx', ky')$ $(k \neq 0)$, at 2. Suppose that the curve C is interior to the region R for $\tau \geqq \tau_2$. Then it follows from the hypotheses of the Theorem that

$$X'_2 F_{x'}(x, x') + Y'_2 F_{y'}(x, x') > 0$$

at 2. Consider the function [10]

$$W(x, y) \equiv I(E_{44'} + \mathcal{E}_{4'2''}),$$

[10] Compare Bolza, *Vorlesungen*, pp. 405-07.

and define

$$S(\tau_{2''}) \equiv W(x_{2''}, y_{2''}).$$

It can be shown that

$$S'(\tau_{2''}) = X'_2 F_{x'}(x, \phi') + Y'_2 F_{y'}(x, \phi')$$

at $2''$, which is positive near 2. It is easily seen that

$$I(\mathcal{E}_{42'}) - I(E_{42}) = S(\tau_{2'}) - S(\tau_2)$$

if the point 4 precedes 2 along E_{12}. Therefore the inequality (13) holds in this case if $\mathfrak{F}_2$ is sufficiently restricted. In the same way we can show that this inequality holds if C is interior to R for $\tau \leqq \tau_2$.

If $D(x', X'_2) = 0$ at 2 and C lies beyond the normal to E_{12} at 2 in such a way that the point 4 lies beyond 2 along E_{56}, then the inequality (11) becomes

$$I(V_{1'2'}) > I(\mathcal{E}_{1'3} + E_{32}) + I(E_{24} + \mathcal{E}_{42'}). \tag{14}$$

An argument similar to the one just given can be used to show that

$$I(E_{24} + \mathcal{E}_{42'}) > 0, \tag{15}$$

in this case if the field $\mathfrak{F}_2$ is sufficiently restricted. From the inequalities (11) and (13) or from (14) and (15), it follows that

$$I(V_{1'2'}) > I(\mathcal{E}_{1'3} + E_{32}). \tag{16}$$

Steps perfectly analogous to those just completed, but making use of the field $\mathfrak{F}_1$ defined by the family of extremals $(\mathcal{E})$ preceding their points of tangency with E_{56}, can be taken to prove that the inequality (12) is true if $\mathfrak{F}_1$ is sufficiently restricted and the point 3 follows 1 along E_{56}. If $D(x', X'_1) = 0$ at 1 and the curve D precedes the normal to E_{12} at 1 in such a way that 3 precedes 1 on E_{56}, then the inequality (16) becomes

$$I(V_{1'2'}) > I(\mathcal{E}_{1'3} + E_{31} + E_{12}), \tag{17}$$

and it can be shown as in the analogous case at 2 that

$$I(\mathcal{E}_{1'3} + E_{31}) > 0. \tag{18}$$

From the inequalities (12) and (16) or from (17) and (18), it follows that

$$I(V_{1'2'}) > I(E_{12}).$$

This completes the proof of Theorem 2 under the assumption that there is some point 8 on E_{12} at which $F(x, x') \neq 0$. If E_{12} is a solution of the differential equation $F(x, x') = 0$, that is if F is identically zero along E_{12}, one may replace the field $\mathfrak{F}_3$ used in the above proof by one determined in a neighborhood of 8 by the one parameter family of extremals through the point 8, for example.

It should be pointed out that if $F(x, x') = 0$ and not both $F_{x'}$ and $F_{y'}$ are zero at 1 or 2, then Theorem 2 may be vacuous as follows from the geometric interpretation of the conditions of the Theorem already given. For example, if this is true at the point 2, then the curve C is not tangent to E_{12} and $\tilde{T}(x, x', X'_2)$ is positive or is negative for all directions $(X'_2, Y'_2) \neq (kx', ky')$ $(k \neq 0)$. In the latter case there is no direction (X'_2, Y'_2) for which the conditions of the Theorem are satisfied.

The next four theorems have to do with the case where $T = 0$ along E_{12}, that is when E_{12} is an arc of an extremal. Therefore, each theorem will involve the Jacobi condition for the arc E_{12}. By means of the methods used by the author [11] and those of Bliss [12] for two variable end-points, one is able to derive the Jacobi condition for each of the cases. These proofs do not require that the enveloping curve have a regressive branch at its point of contact with the minimizing curve. We shall state these conditions strengthened in the usual way.

(I) *E_{12} contains no pair of conjugate points.*

If the curve D is transversal to E_{12} at the point 2, then

(II) *E_{12} does not contain the critical point of the curve D.*

If the curve C is transversal to E_{12} at the point 2, then

(III) *E_{12} does not contain the critical point of the curve C.*

If the curves D and C are transversal to E_{12} at the points 1 and 2 respectively, then

(IV) *The critical point d of the curve D and the critical point c of the curve C lie in the order $12cd$ or some cyclic order with all points distinct.*

---

[11] Mancill, "The Jacobi condition for unilateral variations," *Duke Journal of Mathematics*, vol. 6 (1940), pp. 341-44.

[12] *Loc. cit.*, p. 74.

The next theorem is analogous to Theorem 2 for the case where E_{12} is an arc of an extremal.

THEOREM 3. *Suppose that E_{12} is an extremal arc with the properties:*

1) *condition* (I) *is satisfied by E_{12};*

2) *if $D(x', X'_1) \neq 0$ at 1, then $\bar{T}(x, x', X'_1) < 0$ there; or if $D(x', X'_1) = 0$ at 1, then*

$$F(x, x') > 0 \text{ at } 1 \text{ if } D \text{ precedes the normal to } E_{12} \text{ at } 1,$$
$$F(x, x') < 0 \text{ at } 1 \text{ if } D \text{ lies beyond the normal to } E_{12} \text{ at } 1;$$

3) *if $D(x', X'_2) \neq 0$ at 2, then $\bar{T}(x, x', X'_2) > 0$ there; or if $D(x', X'_2) = 0$ at 2, then*

$$F(x, x') < 0 \text{ at } 2 \text{ if } C \text{ precedes the normal to } E_{12} \text{ at } 2,$$
$$F(x, x') > 0 \text{ at } 2 \text{ if } C \text{ lies beyond the normal to } E_{12} \text{ at } 2.$$

Then the arc E_{12} minimizes the integral I in the class of all admissible curves sufficiently near E_{12} joining the curves D and C.

There exists a one parameter family of extremals

$$x = \phi(t, a), \qquad y = \psi(t, a),$$

through a point 0 preceding the point 1 on E_{12}, containing E_{12} for $a = a_0$, and which forms a field $\mathfrak{F}$ determined by the parameter values

$$a_0 \leqq a \leqq a_0 + e, \qquad t_1 - e \leqq t \leqq t_2 + e,$$

for e sufficiently small. This field lies interior to the region R of admissible curves. If we assume that the curve D is inferior to R for $\tau \geqq \tau_1$, and that C is interior to R for $\tau \geqq \tau_2$, then it follows from the hypotheses of the Theorem that

$$X'_1 F_{x'}(x, x') + Y'_1 F_{y'}(x, x') < 0$$

and

$$X'_2 F_{x'}(x, x') + Y'_2 F_{y'}(x, x') > 0$$

at 1 and 2 respectively. Let $V_{1'2'}$ be any admissible comparison curve in $\mathfrak{F}$ joining the points 1′ and 2′ respectively. Then we have

$$(19) \qquad I(V_{1'2'}) - I(E_{12}) = I(V_{1'2'}) - I^*(E_{12})$$
$$= I(V_{1'2'}) - I^*(V_{1'2'}) + I^*(C_{22'}) - I^*(D_{11'})$$

because of the invariance of the integral I^*. But the right member of (19) is positive since our problem is regular and

$$I^*(D_{11'}) < 0$$

and

$$I^*(C_{22'}) > 0,$$

if the field $\mathfrak{F}$ is sufficiently restricted. Obviously the same conclusion can be reached if the positive direction along either D or C is reversed. Therefore, Theorem 3 is proved.

THEOREM 4. *Suppose that E_{12} is an extremal arc with the properties:*

1) $\bar{T}(x, x', X'_2) = 0$ *and* $F(x, x') \neq 0$ *at the point* 2;

2) *condition* (II) *is satisfied by* E_{12};

3) *same as in Theorem* 3.

Then the arc E_{12} minimizes the integral I in the class of all admissible curves sufficiently near E_{12} joining the curves D and C.

The one parameter family of extremals cut by D transversally forms a field in R in which we may apply the same proof as in Theorem 3, since C is not transversal to E_{12} at the point 2. In the same way we can prove the next theorem.

THEOREM 5. *Suppose that E_{12} is an extremal arc with the properties:*

1) $\bar{T}(x, x', X'_2) = 0$ *and* $F(x, x') \neq 0$ *at the point* 2;

2) *same as in Theorem* 3;

3) *condition* (III) *is satisfied by* E_{12}.

Then the arc E_{12} minimizes the integral I in the class of all admissible curves sufficiently near E_{12} joining the curves D and C.

In the next two theorems we shall treat the case analogous to the problem of free variations with both end-points variable,[13] that is where E_{12} is an extremal and both D and C are transversal to E_{12}. The one parameter family of extremals

$$x = \phi(t, a), \qquad y = \psi(t, a), \tag{20}$$

cut by D transversally contains E_{12} for $a = a_0$. If this family forms a field about E_{12}, we shall denote by C' the transversal of this field through the point 2, which in this case will be tangent to C at 2.

THEOREM 6. *Suppose that E_{12} is an extremal arc with the properties:*

1) $\bar{T}(x, x', X'_1) = 0$ *and* $F(x, x') \neq 0$ *at the point* 1;

2) *condition* (II) *is satisfied by* E_{12};

3) $\bar{T}(x, x', X'_2) = 0$ *and* $F(x, x') \neq 0$ *at the point* 2;

4) $F(x, x') > 0$ *at* 2 *and* C' *coincides with or lies on the same side of* C *as does* D *near* 2 *in* R; *or* $F(x, x') < 0$ *at* 2 *and* C' *coincides with* C *or lies on the side of* C *opposite to* D *near* 2 *in* R.

Then E_{12} minimizes the integral I in the class of all admissible curves joining the curves D and C and sufficiently near E_{12}.

Under the hypotheses of the theorem the family of extremals (20) forms a field about E_{12}. Also there is a unique curve C' through 2 which cuts each of the extremals of the field transversally. Let $V_{1'2'}$ denote any admissible comparison curve in the field intersecting the curve D and C in $1'$ and $2'$ respectively. Suppose that the first part of hypothesis 4) of the theorem is satisfied. Then the usual proof shows that

$$I(V_{1'2'}) > I(\mathcal{E}_{52'}), \tag{21}$$

where $\mathcal{E}_{52'}$ is the particular member of the family (20) through the point 5 on D and the point $2'$ on C. Let the extremal $\mathcal{E}_{52'}$ meet C' in the point 3, and we have

$$I(\mathcal{E}_{52'}) = I(\mathcal{E}_{53} + \mathcal{E}_{32'}) = I(E_{12} + \mathcal{E}_{32'}),$$

since D and C' are both transversals of the field. Since $F(x, x') > 0$ at 2,

[13] For the problem in non-parametric form see, Bliss, *loc. cit.*, pp. 75-80.

it will remain positive along each member of the family (20) in a sufficiently small neighborhood of 2. Therefore, if the field is sufficiently restricted, we have

$$I(\mathcal{E}_{32'}) \geqq 0.$$

Consequently the theorem is proved for the case when C' precedes C or coincides with C along the extremals (20) and $F(x, x') > 0$ at 2.

Now suppose that the second part of hypothesis 4) of the theorem is satisfied. In this case $F(x, x') < 0$ at 2 and

$$I(\mathcal{E}_{2'3}) < 0$$

if the field is sufficiently restricted. Therefore, it follows from the inequality (21) that

$$I(V_{1'2'}) > I(\mathcal{E}_{32'} + \mathcal{E}_{2'3}) = I(E_{12}),$$

since D and C are transversals of the field.

THEOREM 7. *Suppose that E_{12} is an extremal arc with the properties:*

1) $\bar{T}(x, x', X'_1) = 0$ *and* $F(x, x') \neq 0$ *at the point* 1;

2) $\bar{T}(x, x', X'_2) = 0$ *and* $F(x, x') \neq 0$ *at the point* 2;

3) *condition* (IV) *is satisfied along* E_{12};

Then E_{12} minimizes the integral I in the class of all admissible curves joining the curves D and C and sufficiently near E_{12}.

Again the family of extremals (20) forms a field about E_{12}. It has been shown in general that the position of the critical point c of the curve C depends only upon the curvature of C at its intersection 2 with E_{12}. The radius of curvature r of C at 2 is thus a known function of the parameter t of the critical point c. This function has a non-zero derivative dr/dt whose sign is that of

$$F(x, x') \sin (\theta_E - \theta_C)/F_1(x, x')$$

at 2, where θ_E and θ_C are the angles that E_{12} and C respectively make with the x-axis.[14] Now suppose we consider the case where the order of points in condition (IV) is $12cd$. We shall break the proof into two cases:

---

[14] Bliss, "The second variation of a definite integral when one end-point is variable," *Transactions of the American Mathematical Society*, vol. 3 (1902), p. 139. There seems to be an error in sign in the formula for dr/dt on this page.

1) $F(x, x') > 0$ at the point 2:

If C is directed toward the left of E_{12}, that is toward the interior of R at 2, then the derivative $dr/dt < 0$ and $r_C > r_{C'}$, since c precedes d and d is the critical point of C'. But in this case r_C and $r_{C'}$ are negative, and therefore C' lies to the left of C, that is on the same side of C as D. If C is directed toward the right of E_{12}, then the derivative $dr/dt > 0$ and $r_C < r_{C'}$ since c precedes d. But in this case r_C and $r_{C'}$ are positive, and therefore C' lies on the same side of C as D. Thus, Theorem 6 applies in this case.

2) $F(x, x') < 0$ at the point 2:

If C is directed toward the left of E_{12}, the derivative $dr/dt > 0$ and $r_C < r_{C'}$ since c precedes d. But in this case r_C and $r_{C'}$ are positive and C' lies on the side of C opposite to D. If C is directed toward the right of E_{12}, then $dr/dt < 0$ and $r_C > r_{C'}$ since c precedes d. But in this case r_C and $r_{C'}$ are negative and C' lies on the side of C opposite to D. Therefore, Theorem 6 applies.

A similar proof can be made for each of the cases determined by the order of the points 1, 2, c, and d. If the points c and d coincide, then E_{12} may or may not minimize the integral I, but Theorem 6 furnishes a set of sufficient conditions in this case.

The treatment of the case where E_{12} is an extremal and the curves D and C are transversal to E_{12} just completed is the same for free variations. It is the opinion of the author that the treatment given here is simpler than that given by Bliss to which we have already referred.

UNIVERSITY OF ALABAMA.

ON SOME TRIANGULAR SUMMABILITY METHODS.*

By JOSHUA BARLAZ.

1. Let there be given a sequence-to-function transformation

$$\Psi(x) = \sum_{\nu=0}^{\infty} s_\nu \psi_\nu(x), \tag{1.1}$$

where $\{\psi_\nu(x)\}$ is a sequence of functions defined on a point set $\mathcal{E}$ having a limit point ξ not belonging to the set; then the generalized limit of the sequence $\{s_\nu\}$ is defined by

$$\lim_{x\to\xi} \Psi(x) = s = \text{gen lim } s_\nu$$

if the limit on the left exists. In a recent paper,[1] Otto Szász considered the question of selecting a sequence of points $\{x_n\}$ belonging to the point set $\mathcal{E}$, $x_n \to \xi$ as $n \to \infty$, and associating with it the sequence-to-sequence triangular transformation

$$B_n(x_n) = \sum_{\nu=0}^{n} s_\nu \psi_\nu(x_n), \qquad n = 0, 1, 2, \cdots, \tag{1.2}$$

where the generalized limit of the sequence $\{s_\nu\}$ is now defined by

$$\lim_{n\to\infty} B_n(x_n) = s = \text{gen lim } s_\nu.$$

It is to be noted that $B_n(x)$ is defined by taking the first $n+1$ terms of the series defining the function $\Psi(x)$; the summability method is then constructed with the sequence $\{x_n\}$.

As set forth in the aforementioned paper of Szász,[2] the regularity of either method (1.1) or (1.2) does not imply the regularity of the other method. On the other hand, if the first method is boundedness preserving, then the associated triangular methods are also boundedness preserving for all sequences $\{x_n\}$, $x_n \to \xi$. The conditions of regularity for the transformation (1.2) are

$$\lim_{n\to\infty} \psi_\nu(x_n) = 0, \qquad \nu = 0, 1, 2, \cdots, \tag{1.3}$$

* Received February 27, 1946. Presented to the American Mathematical Society, September 15, 1945. The author wishes to thank Professor Otto Szász for his interest in the preparation of this paper.

[1] The superscripts refer to the notes and references at the end of the paper.

$$\sum_{\nu=0}^{n} |\psi_\nu(x_n)| < K, \ K \text{ independent of } n, \tag{1.4}$$

$$\lim_{n\to\infty} \sum_{\nu=0}^{n} \psi_\nu(x_n) = 1. \tag{1.5}$$

The purpose of this paper is to consider the triangular matrix transformation which results when $\psi_\nu(x)$ is taken equal to $e^{-x}x^\nu/\nu!$; that is, when the generating process is applied to Borel's exponential method. The generalized limit of the sequence $\{s_\nu\}$ is now defined by

$$\text{gen lim } s_\nu = \lim_{n\to\infty} e^{-x_n} \sum_{\nu=0}^{n} (s_\nu x_n{}^\nu/\nu!), \qquad x_n \to \infty. \tag{1.6}$$

2. The first problem considered for this summability method is the determination of necessary and sufficient conditions on the sequence of points $\{x_n\}$, $x_n \to \infty$, so that the regularity conditions (1.3), (1.4), and (1.5) are satisfied. It is easily seen that the conditions (1.3) and (1.4) are here fulfilled with every sequence $\{x_n\}$, so long as $x_n \to \infty$. Therefore, in order to insure regularity, it remains to choose the sequence $\{x_n\}$ so that

$$g_n(x_n) \equiv g_n = e^{-x_n} \sum_{\nu=0}^{n} x_n{}^\nu/\nu! \to 1 \text{ as } n \to \infty. \tag{2.1}$$

In determining the sequences $\{x_n\}$ satisfying (2.1), we shall use the following limit due to Gauss:[3]

$$\text{(A)} \qquad \lim_{x\to\infty} e^{-x} \sum_{\nu < x+\omega\sqrt{x}} x^\nu/\nu! = 1/\sqrt{2\pi} \int_{-\infty}^{\omega} e^{-t^2/2}dt \equiv p(\omega).$$

The function $p(\omega)$ has the following properties:

(B)
1. $p(-\infty) = \lim_{\omega\to-\infty} p(\omega) = 0$,
2. $p(0) = 1/2$,
3. $p(\infty) = \lim_{\omega\to\infty} p(\omega) = 1$,
4. $p(\omega)$ increases continuously as ω increases.

THEOREM 2.1. *A necessary and sufficient condition that*

$$g_n(x_n) \to \lambda \text{ as } n \to \infty, \qquad 0 \le \lambda \le 1 \tag{2.2}$$

is that

$$\lim_{n\to\infty} (x_n - n)/\sqrt{n} = -\rho, \qquad x_n \to \infty \tag{2.3}$$

where $p(\rho) = \lambda$, *i. e.*, $\rho = p_{-1}(\lambda)$.

Proof. (a) Sufficiency: We consider first the case when ρ is finite. It follows from (2.3) that for every positive ϵ, however small, there exists an $N_0 = N_0(\epsilon)$ such that

$$n - (\rho + \epsilon)\sqrt{n} \leq x_n \leq n - (\rho - \epsilon)\sqrt{n}, \qquad n > N_0.$$

Transforming the above inequalities, we have

$$(\sqrt{n} - (\rho + \epsilon)/2)^2 - (\rho + \epsilon)^2/4 \leq x_n \leq (\sqrt{n} - (\rho - \epsilon)/2)^2 - (\rho - \epsilon)^2/4.$$

Choosing n so large that $\sqrt{n} > (\rho + \epsilon)/2$, we obtain

$$\sqrt{x_n + (\rho - \epsilon)^2/4} + (\rho - \epsilon)/2 \leq \sqrt{n} \leq \sqrt{x_n + (\rho + \epsilon)^2/4} + (\rho + \epsilon)/2$$

$$x_n + (\rho - \epsilon)\sqrt{x_n + (\rho - \epsilon)^2/4} + (\rho - \epsilon)^2/2$$

$$\leq n \leq x_n + (\rho + \epsilon)\sqrt{x_n + (\rho + \epsilon)^2/4} + (\rho + \epsilon)^2/2.$$

Inasmuch as $x_n \to \infty$, we have for n large enough

$$x_n + (\rho - 2\epsilon)\sqrt{x_n} \leq n \leq x_n + (\rho + 2\epsilon)\sqrt{x_n}.$$

From this it immediately follows that

$$(2.4) \qquad e^{-x_n} \sum_{\nu \leq x_n + (\rho - 2\epsilon)\sqrt{x_n}} x_n^{\nu}/\nu! \leq g_n \leq e^{-x_n} \sum_{\nu \leq x_n + (\rho + 2\epsilon)\sqrt{x_n}} x_n^{\nu}/\nu!;$$

letting n become infinite, we get from (A)

$$p(\rho - 2\epsilon) \leq \underline{\lim}\, g_n \leq \overline{\lim}\, g_n \leq p(\rho + 2\epsilon).$$

Inasmuch as ϵ may be chosen arbitrarily small, and $p(\omega)$ is continuous, we have

$$\lim g_n(x_n) = p(\rho) = \lambda.$$

If ρ is positive infinity, the condition (2.3) leads to the inequality

$$n \geqq x_n + K\sqrt{x_n},$$

K arbitrarily large, n large enough depending on K. This implies that

$$p(K) \leq \underline{\lim}\, g_n \leq \overline{\lim}\, g_n \leq 1,$$

whence it follows that

$$\lim g_n = p(\infty) = 1.$$

If ρ is negative infinity, we obtain similarly

$$n \leq x_n - K\sqrt{x_n},$$

and consequently $0 \leq \underline{\lim}\, g_n \leq \overline{\lim}\, g_n \leq p(-K)$ which yields $\lim g_n = p(-\infty) = 0$. This completes the proof of the sufficiency.

(b) Necessity: The necessity of (2.3) for (2.2) follows easily from (a), the proof of the sufficiency; for, by (a), the sequence $\{(x_n - n)/\sqrt{n}\}$ could not have more than one limit point, finite or infinite, in order that g_n have a limit. Again by (a), the one limit point must be $-\rho$.

THEOREM 2.2. *A necessary and sufficient condition that the summability method defined by* (1.6) *be regular is that*

$$\lim_{n\to\infty} (x_n - n)/\sqrt{n} = -\infty, \qquad x_n \to \infty \text{ as } n \to \infty. \tag{2.5}$$

This theorem follows immediately from the preceding one and the fact that (1.3) and (1.4) are satisfied when $x_n \to \infty$.

It is easy to see that we can derive from Theorem 2.1 the following result: Consider the summability method defined by the relation

$$\text{gen lim } s_\nu = \lim_{n\to\infty} e^{-x_n}/\lambda \sum_{\nu=0}^{n} s_\nu x_n{}^\nu/\nu!, \qquad x_n \to \infty, \quad 0 < \lambda \leq 1. \tag{2.6}$$

A necessary and sufficient condition for the regularity of the method is (2.3). Theorem 2.2 is the special case $\lambda = 1$, $\rho = \infty$.

The transform (1.6) clearly is always boundedness preserving. If $\lim (x_n - n)n^{-\frac{1}{2}}$ exists, the transform is convergence preserving. Furthermore, if $\lim (x_n - n)n^{-\frac{1}{2}} = \infty$, the method sums every bounded sequence to the value zero.

3. The first application that we shall discuss for the transform (1.6) is to analytic functions of a complex variable defined by a power series with a finite radius of convergence. It is the purpose of this section to show that, with certain conditions on the sequence of points $\{x_n\}$, the summability method defined by (1.6) yields the analytic continuation of the function within the Borel polygon.[4,5]

Let there be given a function defined by a power series in a positive circle of convergence:

$$f(z) = \sum_{\nu=0}^{\infty} a_\nu z^\nu, \qquad |z| < r. \tag{3.1}$$

Let

$$s_n(z) = \sum_{\nu=0}^{n} a^\nu z_\nu, \qquad n = 0, 1, 2, \cdots, \tag{3.2}$$

$$F(z, x_n) \equiv F_n(z) = e^{-x_n} \sum_{\nu=0}^{n} s_\nu(z) x_n{}^\nu/\nu!. \tag{3.3}$$

Then it is clear that (3.3) gives the application of (1.6) to the sequence of partial sums defined in (3.2). When $\{x_n\}$ satisfies the condition (2.5), Theorem 2.2 implies that $F_n(z) \to f(z)$ as $n \to \infty$ for $|z| < r$. As has been anticipated, the object is to determine additional conditions on $\{x_n\}$ so that $F_n(z) \to f(z)$ as $n \to \infty$ in a larger region for z, namely the Borel polygon for $f(z)$. We first proceed to set up several formulae relating $F_n(z)$ to $f(z)$ more directly than is done by the defining formula (3.3).

Having started with the function $f(z)$, defined by (3.1), let Γ be any closed Jordan curve containing the origin, such that $f(z)$ is regular interior to and on Γ. Then from the formula

$$a_m = 1/2\pi i \int_\Gamma \frac{f(z)}{z^{m+1}}\, dz, \qquad m = 0, 1, 2, \cdots, \tag{3.4}$$

it follows that

$$s_\nu(z) = \sum_{m=0}^{\nu} a_m z^m = \sum_{m=0}^{\nu} z^m/2\pi i \int_\Gamma \frac{f(\zeta)}{\zeta^{m+1}}\, d\zeta,$$

$$s_\nu(z) = 1/2\pi i \int_\Gamma \frac{f(\zeta)}{\zeta} \left\{ \frac{1 - z^{\nu+1}/\zeta^{\nu+1}}{1 - z/\zeta} \right\} d\zeta, \; z \text{ not on } \Gamma.$$

Assume now that z is interior to Γ. We obtain

$$s_\nu(z) = 1/2\pi i \int_\Gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta - 1/2\pi i \int_\Gamma \frac{f(\zeta) \cdot z^{\nu+1}}{(\zeta - z)\zeta^{\nu+1}}\, d\zeta,$$

$$s_\nu(z) = f(z) - 1/2\pi i \int_\Gamma \frac{f(\zeta) \cdot z^{\nu+1}}{(\zeta - z)\zeta^{\nu+1}}\, d\zeta. \tag{3.5}$$

Substituting now the formula (3.5) for $s_\nu(z)$ into (3.3), we get

$$F_n(z) = e^{-x_n} \sum_{\nu=0}^{n} f(z) x_n{}^\nu/\nu! - e^{-x_n}/2\pi i \sum_{\nu=0}^{n} x_n{}^\nu/\nu! \int_\Gamma \frac{f(\zeta) z^{\nu+1}}{(\zeta - z)\zeta^{\nu+1}}\, d\zeta,$$

$$F_n(z) = e^{-x_n} \sum_{\nu=0}^{n} f(z) x_n{}^\nu/\nu! - e^{-x_n}/2\pi i \int_\Gamma \frac{f(\zeta) z}{(\zeta - z)\zeta} \{ \sum_{\nu=0}^{n} x_n{}^\nu z^\nu/\zeta^\nu \nu! \} d\zeta, \tag{3.6}$$

or

$$F_n(z) \equiv f(z) g_n(x_n) - I(x_n, z), \tag{3.6'}$$

where

$$I(x_n, z) \equiv I_n = e^{-x_n}/2\pi i \int_\Gamma \frac{f(\zeta) z}{(\zeta - z)\zeta} \{ \sum_{\nu=0}^{n} x_n{}^\nu z^\nu/\zeta^\nu \nu! \} d\zeta, \tag{3.7}$$

and $g_n(x_n)$ is given in (2.1). Thus the problem has been reduced to the determination of conditions on $\{x_n\}$ so that $g_n \to 1$, $I_n \to 0$ as $n \to \infty$. Clearly $F_n(z) \to f(z)$ under these conditions.

We now proceed to simplify the integral I_n by means of a transformation of the kernel $\sum_{\nu=0}^{n} x_n^{\nu} z^{\nu}/\zeta^{\nu}\nu!$ into an integral. As has become standard in such cases, the starting point for the conversion is the familiar formula $m! = \int_0^{\infty} t^m e^{-t} dt$. We have

$$K(\lambda) = 1/n! \sum_{\nu=0}^{n} \frac{n!\lambda^{\nu}}{\nu!(n-\nu)!} \int_0^{\infty} t^{n-\nu} e^{-t} dt = \sum_{\nu=0}^{n} \lambda^{\nu}/\nu!,$$

(3.8)

$$K(\lambda) = 1/n! \int_0^{\infty} (\lambda + t)^n e^{-t} dt, \qquad n \geq 0.$$

Write $z/\zeta = \omega = \alpha + i\beta$, α, β real. Then

$$K_n \equiv K(x_n z/\zeta) = 1/n! \int_0^{\infty} (t + (x_n z/\zeta))^n e^{-t} dt$$

$$= 1/n! \int_0^{\infty} (t + \alpha x_n + i\beta x_n)^n e^{-t} dt;$$

$$n! K_n = e^{\alpha x_n} \int_{\alpha x_n}^{\infty} (t + i x_n \beta)^n e^{-t} dt$$

$$= e^{x_n(\alpha+i\beta)} \int_{\alpha x_n}^{\infty} (t + i x_n \beta)^n e^{-(t+i x_n \beta)} dt.$$

Considering $t + ix_n\beta$ as a complex variable s, we have

$$e^{-\omega x_n} n! K_n = \int_C e^{-s} s^n ds$$

where C is the straight line path $x_n\alpha + ix_n\beta \to +\infty + ix_n\beta$. Let $C(\rho)$ designate the straight line path $x_n\alpha + ix_n\beta \to \rho + ix_n\beta$. Then

$$e^{-\omega x_n} n! K_n = \lim_{\rho \to \infty} \int_{C(\rho)} e^{-s} s^n ds. \tag{3.9}$$

From the fact that $e^{-s}s^n$ is a regular function of s in the whole complex plane, it follows that

$$\int_{C(\rho)} e^{-s} s^n ds = \left(\int_{C_1} + \int_{C_2} + \int_{C_3} \right) s^n e^{-s} ds \equiv T_1 + T_2 + T_3 \tag{3.10}$$

where C_1, C_2, C_3 are the straight line paths $x_n\omega \to 0$, $0 \to \rho$, $\rho \to \rho + ix_n\beta$ respectively. It is easily seen that

$$| T_3 | \leq | x_n\beta | \, e^{-\rho} \, | \rho + ix_n\beta |^n = o(1) \text{ as } \rho \to \infty, \tag{3.11}$$

$$T_2 = \int_0^\rho e^{-s} s^n ds \to n! \text{ as } \rho \to \infty. \tag{3.12}$$

We now have from (3.9), (3.10), (3.11), (3.12)

$$K_n = \sum_{\nu=0}^{n} x_n^\nu z^\nu / \zeta^\nu \nu! = e^{\omega x_n} + (e^{\omega x_n}/n!) \int_{C_1} s^n e^{-s} ds. \tag{3.13}$$

Combining (3.7) and (3.13), we obtain

$$I_n = 1/2\pi i \int_\Gamma \frac{f(\zeta) z}{(\zeta - z)\zeta} e^{-x_n(1-z/\zeta)} d\zeta + R_n(z), \tag{3.14}$$

where

$$R_n(z) = 1/2\pi i \int_\Gamma \frac{f(\zeta) z}{(\zeta - z)\zeta} e^{-x_n(1-z/\zeta)} \{1/n! \int_{C_1} s^n e^{-s} ds\} d\zeta. \tag{3.15}$$

THEOREM 3.1. *If $x_n = o(n)$ as $n \to \infty$, then $R_n(z) \to 0$ as $n \to \infty$, for all z (not on Γ).*

Proof. From the hypothesis, it follows that for any positive ϵ, however small, $x_n \leq \epsilon n$ for n large enough. Using Stirling's formula: $n! \approx n^n e^{-n} (2\pi)^{1/2}$, we get

$$| e^{-x_n(1-\omega)}/n! \int_{C_1} s^n e^{-s} ds | \leq e^{-x_n(1-a)} e^{|a| x_n} | x_n z/\zeta |^{n+1}/n!$$

$$\leq K \frac{e^{A x_n} B^{n+1} n^{n+1} \epsilon^{n+1}}{n^n e^{-n} (2\pi n)^{1/2}}, \quad n > n_0.$$

Recalling that Γ is a closed Jordan curve containing the origin, we see that K, A, B are absolute constants independent of ζ, ζ on Γ. Hence, if ϵ is chosen small enough, it follows that the right hand member of the above inequality tends to zero uniformly as to ζ, ζ on Γ, as $n \to \infty$. This is sufficient to prove the theorem.[6]

THEOREM 3.2. *If $x_n = o(n)$, $x_n \to \infty$, as $n \to \infty$, then $F_n(z) \to f(z)$ as $n \to \infty$, for all z in the Borel polygon of the function $f(z)$.*

Proof. From the condition $x_n = o(n)$, Theorem 2.1 implies that $g_n \to 1$ as $n \to \infty$. If z satisfies the inequality $\mathcal{R}(z/\zeta) < 1$ for all ζ on Γ, clearly

$$1/2\pi i \int_\Gamma \frac{f(\zeta) z}{(\zeta - z)\zeta} e^{-x_n(1-z/\zeta)} d\zeta \to 0, \qquad n \to \infty. \tag{3.16}$$

It is well known that the condition $\mathcal{R}(z/\zeta) < 1$ for all ζ on Γ implies that,

with the proper choice of the curves Γ, (3.16) can be proved for all z in the polygon.[7] Applying the formulae (3.6′), (3.14), and Theorem 3.1, the proof is rendered complete.

Theorem 3.2 gives a condition under which the triangular summability method defined by (1.6) sums a power series within the Borel polygon of the function represented by this series.[8] However, we can also use the formulae derived on the preceding pages to show that if $x_n = o(n)$, then the transform (1.6) is as powerful as Borel's exponential method of summation when applied to a power series with a non-zero radius of convergence.

To demonstrate the result just mentioned, let us return to formula (3.5). We observe that if z is exterior to Γ, then (3.5) becomes

$$s_\nu(z) = -1/2\pi i \int_\Gamma f(\zeta) z^{\nu+1}/(\zeta - z)\zeta^{\nu+1} d\zeta. \tag{3.17}$$

Using (3.3), (3.5), 3.6′), (3.14), (3.17), we can now write for all z not on Γ

$$F(z, x_n) = \delta(z, \Gamma) f(z) g_n(x_n) - 1/2\pi i \int_\Gamma \frac{f(\zeta) z}{(\zeta - z)\zeta} e^{-x_n(1-z/\zeta)} d\zeta - R_n(z \tag{3.18}$$

where Γ is any closed Jordan curve containing the origin, such that $f(z)$ is regular interior to and on Γ, and

$$\begin{aligned} \delta(z, \Gamma) &= 1, \text{ if } z \text{ is interior to } \Gamma \\ &= 0, \text{ if } z \text{ is exterior to } \Gamma. \end{aligned} \tag{3.19}$$

The equation giving the application of Borel's exponential method to the series in (3.1) is

$$B(z, x) = \delta(z, \Gamma) f(z) - 1/2\pi i \int_\Gamma \frac{f(\zeta) z}{(\zeta - z)\zeta} e^{-x_n(1-z/\zeta)} d\zeta, \quad x \to \infty. \tag{3.20}$$

Combining (3.18) and (3.20), we have at once

$$B(z, x_n) - F(z, x_n) = \delta(z, \Gamma) f(z) [1 - g_n] + R_n(z).$$

Using Theorems 2.1 and 3.1, we have

THEOREM 3.3. *If $x_n = o(n)$, then the transform (1.6) is as powerful as Borel's method in summing power series with a positive radius of convergence.*

4. In this section we consider the application of the transform (1.6) to Fourier series. It is immediately seen from Theorem 3.3 that if $x_n = o(n)$, then the triangular method will sum a Fourier series whenever Borel's method does. However, the proof of Theorem 3.2 and the footnote referred to therein suggest the possibility of less restrictive conditions on the sequence $\{x_n\}$ in the case of Fourier series.

Let $f(\theta)$ be a Lebesgue integrable function of period 2π. Following the usual notation for Fourier series, we have

$$f(\theta) \sim a_0/2 + \sum_1^\infty (a_n \cos n\theta + b_n \sin n\theta) \equiv \sum_0^\infty A_n(\theta). \tag{4.1}$$

$$A_n(\theta) \to 0 \text{ as } n \to \infty, \text{ uniformly as to } \theta. \tag{4.2}$$

$$\pi a_n = \int_0^{2\pi} f(t) \cos ntdt, \qquad \pi b_n = \int_0^{2\pi} f(t) \sin ntdt, \qquad n \geqq 0. \tag{4.3}$$

Writing

$$s_n(\theta) = \sum_{\nu=0}^n A_\nu(\theta), \quad n \geq 0, \quad s_n^*(\theta) = s_n(\theta) - \tfrac{1}{2}A_n(\theta), \quad n > 0, \quad s^*_0 = 0,$$

it follows that

$$s_\nu^*(\theta) = 1/\pi \int_0^\pi \psi(\theta, t) \cot (t/2) \sin \nu t dt, \text{ where} \tag{4.4}$$

$$\psi(\theta, t) = \tfrac{1}{2}\{f(\theta + t) + f(\theta - t)\}. \tag{4.5}$$

From (4.2) and the fact that the method (1.6) is linear and boundedness preserving for all sequences $\{x_n\}$, therefore preserving convergence to zero as well as uniform convergence to zero, it follows that we may use the sequence $\{s_n^*(\theta)\}$ instead of $\{s_n(\theta)\}$ in the formal application of (1.6).

Write

$$P_n \equiv P(\theta, x_n) = e^{-x_n} \sum_{\nu=0}^n s_\nu^*(\theta) x_n^\nu/\nu!, \qquad n = 0, 1, 2, \cdots. \tag{4.6}$$

Then from (4.4)

$$P_n = e^{-x_n}/\pi \int_0^\pi \psi(\theta, t) \cot (t/2) \{ \sum_{\nu=0}^n x_n^\nu \sin \nu t/\nu!\} dt. \tag{4.7}$$

Employing (3.8) and (3.13), we obtain

$$\sum_{\nu=0}^{n} x_n{}^{\nu} \sin \nu t/\nu ! = \mathfrak{I}\{\sum_{\nu=0}^{n} x_n{}^{\nu} e^{i\nu t}/\nu !\} = \mathfrak{I}\{K(x_n e^{it})\}$$

$$= \mathfrak{I}\{\exp(x_n e^{it})\} + \mathfrak{I}\{\exp(x_n e^{it}) \int_{C_1} s^n e^{-s}/n !\, ds\},$$

where C_1 is the straight line path $x_n e^{it} \to 0$. Therefore

$$(4.8) \quad P_n = 1/\pi \int_0^{\pi} \psi(\theta, t) \cot(t/2)\, e^{-x_n(1-\cos t)} \sin(x_n \sin t)\, dt + \Omega_n$$

$$\equiv B(\theta, x_n) + \Omega_n$$

where

$$(4.9) \quad \Omega_n = 1/\pi \int_0^{\pi} \psi(\theta, t) \cot(t/2)\, \mathfrak{I}\{\exp(-x_n + x_n e^{it}) \int_{C_1} s^n e^{-s}/n !\, ds\} dt,$$

and $B(\theta, x)$ is the same as the familiar expression which is obtained when Borel's exponential mean is applied to Fourier series.

THEOREM 4.1. (i) *If*

$$(4.10) \quad x_n \leq n/e + K, \quad n > N, \quad K \text{ a constant},$$

then $\Omega_n = o(1)$ *as* $n \to \infty$, *uniformly as to* θ.

(ii) *If* $\{x_n\}$ *satisfies* (4.10), *then* (1.6) *will sum a Fourier series whenever Borel's mean sums the series.*

Clearly the second part of the theorem is an immediate consequence of the first part. Writing

$$D_n = \exp(-x_n + x_n e^{it}) \int_{C_1} s^n e^{-s}/n !\, ds,$$

we have

$$D_n = \exp(-x_n + x_n e^{it} + i(n+1)t) \int_{x_n}^{0} w^n \exp(-we^{it})/n !\, dw,$$

$$|D_n| \leq \int_0^{x_n} \exp(x_n \cos t - x_n - w \cos t)\, w^n/n !\, dw.$$

Observing that $x_n \cos t - x_n - w \cos t \leq 0$ for $0 \leq w \leq x_n$ and for all t, we have

$$|D_n| \leq x_n{}^{n+1}/(n+1) ! \approx (ex_n)^{n+1}/(n+1)^{n+1}(2\pi)^{\frac{1}{2}}(n+1)^{\frac{1}{2}}.$$

Hence if

$$(4.11) \qquad \{ex_n/(n+1)\}^{n+1} = o((n+1)^{1/2}) \text{ as } n \to \infty,$$

it follows that $D_n = o(1)$ as $n \to \infty$, uniformly as to t. Inasmuch as (4.10) implies (4.11), the theorem is proved.

By virtue of the theorem just demonstrated, we note that the many results concerning the application of Borel's method to Fourier series, which have been derived in the literature, can be transformed into theorems for the triangular method (1.6).[9]

5. We conclude this paper with two results for the transform (1.6), which are known to exist for Borel summability.

THEOREM 5.1. *The summability method of arithmetic means is not included in the method* (1.6) *for any sequence* $\{x_n\}$.

Proof. For certain sequences $\{x_n\}$, the theorem is immediate from the results of the preceding section on Fourier series. However, we use here a general method invoking the Silverman-Toeplitz conditions for regularity. Let

$$t_n = e^{-x_n} \sum_{\nu=0}^{n} s_\nu x_n{}^\nu/\nu!, \qquad \sigma_n = \{s_0 + s_1 + \cdots + s_n\}/(n+1), \quad n \geq 0.$$

Clearly

$$t_n = e^{-x_n} \sum_{\nu=0}^{n} \{(\nu+1)\sigma_\nu - \nu\sigma_{\nu-1}\} x_n{}^\nu/\nu!, \qquad \sigma_{-1} = 0$$

$$(5.1) \qquad t_n = e^{-x_n} \sum_{\nu=0}^{n-1} (\nu+1)\sigma_\nu \{x_n{}^\nu/\nu! - x_n{}^{\nu+1}/(\nu+1)!\}$$

$$+ e^{-x_n} x_n{}^n (n+1)\sigma_n/n! \equiv \sum_{\nu=0}^{n} a_{n\nu}\sigma_\nu,$$

where

$$(5.2) \qquad a_{n\nu} = e^{-x_n}(\nu+1)\{x_n{}^\nu/\nu! - x_n{}^{\nu+1}/(\nu+1)!\},$$

$$\nu \leq n-1, \qquad a_{nn} = e^{-x_n}(n+1)x_n{}^n/n!.$$

In order that the transformation of $\{\sigma_\nu\}$, defined by (5.1), be regular, it is necessary that

$$(a) \qquad \sum_{\nu=0}^{n} a_{n\nu} = e^{-x_n}\{\sum_{\nu=0}^{n-1} (\nu+1)(x_n{}^\nu/\nu! - x_n{}^{\nu+1}/(\nu+1)!) + (n+1)x_n{}^n/n!\}$$

$$= e^{-x_n} \sum_{\nu=0}^{n} x_n{}^\nu/\nu! \to 1, \qquad n \to \infty.$$

(b) $$q_n = \sum_{\nu=0}^{n} |a_{n\nu}| = e^{-x_n}\{\sum_{\nu=0}^{\mu-1}(\nu+1)|x_n^{\nu}/\nu! - x_n^{\nu+1}/(\nu+1)!| + (n+1)x_n^{n}/n!\} = O(1) \text{ as } n \to \infty.$$

Theorem 2.1 shows that (a) is satisfied if $\{x_n\}$ satisfies (2.5). In fact, we shall admit those sequences $\{x_n\}$ satisfying (2.3) with $\rho > -\infty$. Then for some k,

$$x_n < n + k\sqrt{n}, \qquad n \geq 1. \tag{5.3}$$

On the contrary, (b) cannot be fulfilled if $\{x_n\}$ satisfies (5.3). To prove this, we consider two cases.

(i) Suppose $n \leq x_n < n + k\sqrt{n}$. Then

$$q_n = -e^{-x_n}\sum_{\nu=0}^{n-1}(\nu+1)\{x_n^{\nu}/\nu! - x_n^{\nu+1}/(\nu+1)!\} + e^{-x_n}(n+1)x_n^{n}/n!$$

$$q_n = 2e^{-x_n}(n+1)x_n^{n}/n! - e^{-x_n}\sum_{\nu=0}^{n} x_n^{\nu}/\nu! > 2e^{-x_n}nx_n^{n}/n! - 1.$$

It is easy to see that the function $e^{-x}x^n$, $x, n > 0$, attains its maximum when $x = n$ and that it is a monotone decreasing function of x, $x > n$. Therefore

$$q_n > 2e^{-n-kn^{1/2}} \cdot n \cdot (n + kn^{1/2})^n/n! - 1 \equiv 2v_n - 1.$$

Inasmuch as

$$v_n \approx (n/2\pi)^{1/2}\, e^{-kn^{1/2}}(1 + kn^{-1/2})^n \approx (n/2\pi)^{1/2}e^{-k^2/2},$$

it follows that q_n is of the order $n^{1/2}$ for the case being considered.

(ii) Suppose $x_n < n$. Letting $\mu = [x_n]$, we have

$$q_n = e^{-x_n}\{-\sum_{\nu=0}^{\mu-1}(\nu+)\,(x_n^{\nu}/\nu! - x_n^{\nu+1}/(\nu+1)!) + \sum_{\nu=\mu}^{n-1}(\nu+1)\,(x_n^{\nu}/\nu! - x_n^{\nu+1}/(\nu+1)!) + (n+1)x_n^{n}/n!\}$$

$$= e^{-x_n}\{-2\sum_{\nu=0}^{n-1}(\nu+1)\,(x_n^{\nu}/\nu! - x_n^{\nu+1}/(\nu+)!) + \sum_{\nu=0}^{n} x_n^{\nu}/\nu!\}$$

$$= 2e^{-x_n}(\mu+1)x_n^{\mu}/\mu! - 2e^{-x_n}\sum_{\nu=0}^{\mu} x_n^{\nu}/\nu! + e^{-x_n}\sum_{\nu=0}^{n} x_n^{\nu}/\nu!$$

$$> 2e^{-x_n}\mu \cdot x_n^{\mu}/\mu! - 2.$$

Recalling that $\mu \leq x_n < \mu + 1$, and using the monotone nature of the function $e^{-x}x^{\mu}$ for $x > \mu$, we obtain

$$q_n > 2e^{-\mu-1} \cdot \mu \cdot (\mu+1)^{\mu}/\mu! - 2 \equiv 2w_{\mu} - 2.$$

Inasmuch as $w_{\mu} \approx (\mu/2\pi)^{1/2}$, it follows that q_n is of the order $x_n^{1/2}$ for this case.

Combining the results of (i) and (ii), we see that (b) and (5.3) are incompatible. Therefore (a) and (b) cannot be satisfied with the same sequence $\{x_n\}$, thus completing the proof of the theorem.

THEOREM 5.2. *The method* (1.6) *sums the sequence* $s_{\nu} = (-1)^{\nu}$, $\nu = 0, 1, 2, \cdots$, *to the value zero for all* $\{x_n\}$.

Proof. Writing $e^{x_n}t_n = \sum_{\nu=0}^{n} (-1)^{\nu}x_n^{\nu}/\nu! = K(-x_n)$, and applying (3.8), we obtain

$$t_n = e^{-x_n}/n! \int_0^{\infty} (t - x_n)^n e^{-t} dt = e^{-2x_n}/n! \int_{-x_n}^{\infty} u^n e^{-u} du,$$

$$t_n = e^{-2x_n}/n! \int_0^{\infty} u^n e^{-u} du + (-1)^n (e^{-2x_n}/n!) \int_0^{x_n} u^n e^{u} du,$$

$$t_n \equiv e^{-2x_n} + (-1)^n r_n.$$

Using the fact that the maximum of the function $u^n e^{-u}$ occurs at $u = n$, we have

$$r_n = e^{-2x_n}/n! \int_0^{x_n} e^{2u} u^n e^{-u} du \leq e^{-2x_n} n^n e^{-n}/n! \int_0^{x_n} e^{2u} du$$

$$\leq K n^{-1/2} e^{-2x_n} (e^{2x_n} - 1) = o(1), \qquad n, x_n \to \infty.$$

Therefore $t_n = o(1)$, $n \to \infty$, which was to be proved.

Theorem 5.2 shows that the method (1.6) is not included in convergence, for any sequence $\{x_n\}$.

OHIO STATE UNIVERSITY.

NOTES.

---

1. O. Szász, "Some new summability methods with applications," *Annals of Mathematics*, vol. 43 (1942), pp. 69-83.

2. In the paper referred to in note 1, the problem was considered for Abel and related summability methods.
3. See (i) M. Kac, "Note on the partial sums of the exponential series," *Univ. Nac. Tucumán, Revista* A. 3 (1942), pp. 151-153, *Mathematical Reviews*, vol. 4 (1943), p. 194; (ii) J. V. Uspensky, *Introduction to Mathematical Probability*, New York, 1937; (iii) Pólya and Szegö, *Aufgaben und Lehrsatze aus der Analysis*, Berlin, vol. 1 (1925), p. 80.
4. R. G. Cooke, "A note on lower semi-matrices," *Journal of the London Mathematical Society*, vol. 14 (1939), pp. 154-157, has also considered the problem of analytic continuation with such triangular summability methods.
5. For the definition of the Borel polygon, see (i) E. Borel, *Leçons sur les séries divergentes*, Paris, 1928; (ii) *Introduction to the Theory of Divergent Series*, lithoprinted lectures by O. Szász, written by J. Barlaz, Ann Arbor, 1944.
6. We took $x_n = o(n)$ so that $R_n \to 0$ for all z. Clearly $R_n \to 0$ for certain z even if x_n is not $o(n)$.
7. See the books referred to in note 5.
8. Referring back to note 6, we see that it is possible to have analytic continuation even if x_n is not $o(n)$. However, the region would be smaller than the polygon.
9. For literature concerning the application of Borel's method to Fourier series, see (i) G. H. Hardy, "Remarks on some points in the theory of divergent series," *Annals of Mathematics*, vol. 36 (1935), pp. 167-181; (ii) G. H. Hardy and J. E. Littlewood, "Some new convergence criteria for Fourier series," *Annali della Scuola Normali di Pisa*, vol. 3 (1934), pp. 43-62; (iii) C. N. Moore, "On the application of Borel's method to the summation of Fourier series," *Proceedings of the National Academy of Sciences*, vol. 11 (1925), pp. 284-287; (iv) L. Lorch, "The Lebesgue constants for Borel summability," *Duke Mathematical Journal*, vol. 11 (1944), pp. 459-467.

TOPOLOGICAL RINGS.*

By Irving Kaplansky.

1. Introduction. The paper consists of two sections which may, for the most part, be read independently. Part I is devoted mainly to the study of compact rings. We first give some preliminary results, including a summary of Jacobson's [11] theory of the radical in a form convenient for our use, some remarks on "Q-rings" (rings in which elements near zero are quasi-regular), and an abstract theory of boundedness. The structure theorems for compact rings are given in Theorems 14-20. A noteworthy feature is the extent to which compactness serves as a substitute for the classical chain conditions.

In Part II the ring of continuous functions from a topological space X to a topological ring A is studied. When X is totally disconnected, the theory of the maximal, prime, or primitive ideals can be worked out satisfactorily with mild restrictions on A. We also give a generalization of the classical case where A is the ring of real or of complex numbers. The paper concludes with a study of two subrings: the bounded functions and those vanishing at a point.

PART I. The Structure of Compact Rings.

2. Q-rings. By a *topological ring* A we shall mean an (associative) ring which is also a Hausdorff space, such that $a - b$ and ab are continuous functions of a and b. The additional assumption of the continuity of the inverse, whenever it exists, will be considered below.

In a *metric* ring there is associated with every element a real number $|a|$ such that $|0| = 0$, $|a| > 0$ for $a \neq 0$, $|-a| = |a|$, $|a+b| \leqq |a| + |b|$, $|ab| \leqq |a|\,|b|$. If in the last statement equality holds, we say that the ring has a *valuation*. Evidently a ring with valuation can have no divisors of zero.

In A we introduce the operation o defined by $xoy = x + y + xy$. It is readily verified to be an associative operation with 0 acting as an identity

* Received September 18, 1946.

element. In general A does not form a group or even a semigroup under o, for the cancellation law may fail.

An element is said to be right quasi-regular (r. q. r.) if there exists an element y with $xoy = 0$; the latter is a right quasi-inverse of x. Left quasi-inverses and quasi-regularity are similarly defined. As is well known for associative systems, if x has a right quasi-inverse y and a left quasi-inverse z, then $y = z$; we then call y the quasi-inverse of x and say that x is quasi-regular (q. r.).

The following related result can be proved: if x has a *unique* right quasi-inverse y, then y is also a left quasi-inverse. For it follows from the identity

$$xo(yox + y) = (xoy)(2 + x)$$

that $yox + y$ is a right quasi inverse of x. Hence $yox + y = y$, $yox = 0$.

A right ideal is r. q. r. if it consists entirely of r. q. r. elements. The following lemma is useful.

LEMMA 1. *If a is in a* r. q. r. *ideal, and x is* r. q. r., *then $a + x$ is* r. q. r.

Proof. Suppose $xoy = 0$. Since $a + ay$ is in the right ideal generated by a, we have $(a + ay)oz = 0$. The following identity then shows that yoz is a right quasi-inverse of $a + x$:

$$(a + x)oyoz = (a + ay)oz + (xoy)(1 + z). \tag{1}$$

We define the radical R of A to be the (set-theoretic) join of all r. q. r. ideals. It follows readily from Lemma 1 that R is a right ideal. To prove that R is also a left ideal, we must show that $ba \in R$ where $a \in R$, $b \in A$, and for this it will suffice to prove that $ba(c + i)$ is r. q. r. for any integer i and $c \in A$. Let $x = a(c + i)$ and let y be a right quasi-inverse of xb. The identity

$$bxo(-bx - byx) = -b(xboy)x$$

shows that bx is r. q. r.

A ring A whose radical is all of A is said to be a radical ring; one whose radical is 0 is semi-simple.

For brevity we shall call a topological ring a Q_r-ring if its r. q. r. elements form an open set. For rings with unity element this is equivalent to saying that the elements with right inverses form an open set. The following lemma provides a criterion which is often more convenient.

LEMMA 2. *A ring which has a neighborhood of* 0 *consisting of* r. q. r. *elements is a Q_r-ring.*

Proof. Let x be any r. q. r. element and y a right quasi-inverse. By taking a sufficiently small we can be sure that $a+ay$ is r. q. r., say $(a+ay)oz=0$. Then by (1), $a+x$ is r. q. r. Hence x has a neighborhood consisting of r. q. r. elements.

If the left q. r. elements are open we shall call the ring a Q_l-ring, and if the q. r. elements are open, a Q-ring. It follows from Lemma 2 that a ring is a Q-ring if and only if it is both a Q_l-ring and a Q_r-ring. However I do not know of any Q_r-ring which is not a Q-ring.

The force of the hypothesis that a ring is a Q-ring can be assessed in part from the fact (Theorem 21) that a locally compact ring without divisors of zero is a Q-ring, and from the fact that any complete metric ring is a Q-ring. The last assertion follows from Lemma 2 and the following lemma.

LEMMA 3. *In a complete metric ring an element x with $|x|<1$ is* q. r.

Proof. The series $-x+x^2-x^3\cdots+(-x)^n+\cdots$ converges say to y, and we have $xoy=yox=0$.

Thus the condition that a ring be a Q-ring is somewhat related to its completeness. However completeness is neither necessary nor sufficient.

(a) Example of a Q-ring which is not complete: the rationals under the ordinary topology.

(b) Example of a complete (even compact) ring which is not a Q-ring: the Cartesian direct sum of an infinite number of finite fields.

Let x' denote the quasi-inverse of x, whenever this exists. A question of some importance is whether the function $x\to x'$ is continuous. We first note the following result.

LEMMA 4. *In a Q-ring the quasi-inverse is continuous wherever defined if it is continuous at* 0.

Proof. Let U be any neighborhood of x'. There exists a neighborhood V of 0 such that $z\in V$ implies $x'oz\in U$, and a neighborhood W of 0 such that $a\in W$ implies that $a+ax'$ has a quasi-inverse z in V. Then (1), with $y=x'$, shows that the quasi-inverse of $a+x$ is $x'oz$, and the latter lies in U.

LEMMA 5. *The quasi-inverse is continuous in a complete metric ring.*

Proof. By Lemmas 3 and 4, it suffices to prove continuity at 0. Since $x'=\Sigma(-x)^n$, we have that $|x|<\mathrm{Min}(1/2,\epsilon/2)$ implies $|x'|<\epsilon$.

We shall show below (Theorem 22) that the quasi-inverse is also continuous in a locally compact ring without divisors of 0. However to encompass

more general cases we shall in general *assume the continuity of the quasi-inverse as an additional hypothesis.* It is to be noted that in case the ring has a unity element, this axiom coincides with the assumption of the continuity of the ordinary inverse, which has often been added to the axioms for a topological ring.

With the aid of the foregoing axiom we can prove the following theorem.

THEOREM 1. *If A is a Q-ring with continuous quasi-inverse, and A_n is the ring of all n by n matrices with elements in A, then A_n is a Q-ring.*

Proof. It is understood that the topology in A_n is that of the Cartesian product, i. e., if U is a neighborhood of 0 in A, then a neighborhood of 0 in A_n consists of all matrices (a_{ij}) with $a_{ij} \in U$.

For a matrix (a_{ij}) with elements suitably near 0, we then seek a matrix (b_{ij}) with

$$a_{ij} + b_{ij} + \sum_{k=1}^{n} a_{ik} b_{kj} = 0 \qquad (i, j = 1, \cdots, n). \tag{2}$$

Keeping j temporarily fixed, this gives n equations for the n unknowns $b_{1j}, \cdots, b_{nj}$. It is convenient to establish the existence of a solution of these equations by the following lemma.

LEMMA 6. *If c_{ij}, d_i $(i, j = 1, \cdots, n)$ are elements of a Q-ring within a sufficiently small neighborhood of 0, then the equations*

$$x_i + \sum_{j=1}^{n} c_{ij} x_j = d_i \qquad (i = 1, \cdots, n) \tag{3}$$

have a solution.

Proof. Let c' be the quasi-inverse of c_{11} (it exists if c_{11} is near enough to 0). We have $c' o (x_1 + c_{11} x_1) = c' + x_1$. Then on applying $c' o$ on the left of the first of the equations (3), we get an equation which is explicitly solved for x_1. We substitute for x_1 in the remaining $n - 1$ equations and obtain a system of equations of the same form as (3), with coefficients which are polynomials in the c_{ij}, d_i, and c'. By induction we may assume that these $n - 1$ equations can be solved, provided their coefficients are sufficiently near 0, and this can be assured by taking the c's and d's sufficiently near 0. By virtue of our assumption of the continuity of the quasi-inverse, c' will be near 0 along with c_{11}.

Since the equations (2) are of the form (3), it follows from Lemma 6

that (2) can be solved for b_{ij} provided (a_{ij}) is near 0, and then the matrix (a_{ij}) has the right quasi-inverse (b_{ij}). Hence A_n is a Q_r-ring. Similarly A_n may be proved to be a Q_l-ring, and therefore it is a Q-ring.

Remark. If A is commutative, we can rearrange the above proof so as to avoid the assumption of the continuity of the quasi-inverse. The ordinary algebraic method of elimination would be available: multiply the first of equations (3) by c_{i1}, the i-th by $1 + c_{11}$, and subtract $(i = 2, \cdots, n)$.

3. The radical. We shall consider in this section questions related to the nature of the radical and in particular to whether it is closed. First we prove the following lemma.[1]

LEMMA 7. *If e is an idempotent in the Q_r-ring A, and B is a right ideal dense in eA, then $B = eA$.*

Proof. B contains elements $e + x$ with x arbitrarily small, say so small that $xeoy = 0$. Left-multiplying by e and using $ex = x$, we obtain $xeoey = 0$. Then

$$(e + xe)(e + y) = e + xeoey = e.$$

Since $e + xe = (e + x)e \in B$, we have $e \in B$ and $B = eA$.

Let M be a maximal right ideal in a Q_r-ring A which has a left unity element. Then the closure of M is either M or A. But in the latter case Lemma 7 implies that $M = A$. The same argument works for a maximal ideal.[2] We therefore obtain the following theorem.

THEOREM 2. *In a Q_r-ring with left unity element the maximal (right) ideals are closed.*

Since [3] in a ring with left unity element the radical R is the intersection of the maximal right ideals, Theorem 2 implies further that R is closed. This may be proved for Q_r-rings without unity element.

THEOREM 3. *In a Q_r-ring the radical is closed.*

Proof. Let y be an element in the closure $\bar{R}$ of R. We can write $y = a + x$ with $a \in R$ and x r. q. r. By Lemma 1, y is r. q. r. Hence $\bar{R}$ consists of r. q. r. elements, $\bar{R} \subseteq R$, $\bar{R} = R$.

---

[1] Prof. Ambrose informs me that Lemma 7 fills in a slight gap in the proof of Theorem 4. 3 of [1].

[2] "Ideal" will always mean two-sided ideal.

[3] [11, Th. 18, Cor. 2]. The proof given there actually makes use only of a left unity element.

THEOREM 4. *If the* r. q. r. *elements are closed, then the radical is closed.*

Proof. Again $\bar{R}$ will consist of r. q. r. elements, whence $\bar{R} = R$.

Theorems 3 and 4 provide useful classes of rings with closed radicals. We shall now give an example of a ring with a radical which is not closed.[4]

Let C be the set of all power series

$$\alpha = a_0 + a_1 x + \cdots + a_n x^n + \cdots$$

with real coefficients, such that $\Sigma \mid a_i \mid /i!$ converges. This sum is to be the norm of α, written $\| \alpha \|$. Elements α and β are added and multiplied in the usual manner for formal power series. It is readily verified that $\alpha + \beta$ and $\alpha\beta$ are again elements of C and that we have $\| \alpha + \beta \| \leqq \| \alpha \| + \| \beta \|$, $\| \alpha\beta \| \leqq \| \alpha \| \, \| \beta \|$, and $\| a\alpha \| = \mid a \mid \, \| \alpha \|$ for a real. Next, C is complete. For let $\alpha_i = \Sigma a_{ij} x^j$ $(i = 1, 2, \cdots)$ be a sequence of power series with $\| \alpha_m - \alpha_n \| \to 0$ as $m, n \to \infty$. We have $\mid a_{mj} - a_{nj} \mid /j! \leqq \| \alpha_m - \alpha_n \|$, and hence for fixed j the sequence $a_{1j}, a_{2j}, \cdots$ converges say to a_j. The norms $\| \alpha_i \|$ are bounded, say by M. Then

$$\sum_{j=0}^{n} \mid a_j \mid /j! = \lim_{i \to \infty} \sum_{j=0}^{n} \mid a_{ij} \mid /j!$$

must also be bounded by M. Hence $\Sigma \mid a_j \mid /j!$ converges, and the series $\alpha = \Sigma a_j x^j$ is in C. Given ϵ, suppose that $\| \alpha_i - \alpha_p \| < \epsilon$ for $i \geqq p$. Then $\| \alpha - \alpha_p \| \leqq \epsilon$. For if not, then for some n

$$\epsilon < \sum_{j=0}^{n} \mid a_j - a_{pj} \mid /j! = \lim_{i \to \infty} \sum_{j=0}^{n} \mid a_{ij} - a_{pj} \mid /j!,$$

which is impossible.

The ring C is a real normed ring in the sense of Gelfand [7]. We next show that its radical is the ideal (x), i. e., the set of all polynomials with constant term zero. Given any such power series $\alpha = a_1 x + a_2 x^2 + \cdots = x\beta$, with $\| \beta \| = a$, we have $\| \alpha^n \| \leqq \| x^n \| \, \| \beta^n \| \leqq \mid a^n \mid /n!$. The series $\gamma = -\alpha + \alpha^2 - \alpha^3 \cdots$ converges since it is dominated by the series for e^a, and we have $\gamma \circ \alpha = \alpha \circ \gamma = 0$. Thus every element in the ideal (x) is q. r., so that the radical contains (x). Moreover it is evident that no power series with a constant term different from zero can be in the radical. Hence the radical is (x). (We could also verify that $\| \alpha^n \|^{1/n} \to 0$ and use Gelfand's characterization of the radical.)

---

[4] The ring C was constructed in conversation with H. Rubin.

Now let D be the set of all sequences $X = (\alpha_1, \alpha_2, \alpha_3, \cdots)$ of elements of C with the property that $|| \alpha_n ||$ increases no faster than a polynomial in n. Formally: for some k (depending on X) and all n we have $|| \alpha_n || < n^k$. Elements of D are to be added and multiplied component-wise, and the topology is that of point-wise convergence. It is evident that D is a topological ring.

We define a sequence Y_n of elements of D by

$$Y_n = (x, 2x, \cdots, nx, 0, 0, \cdots).$$

These elements approach the element $Y \varepsilon D$ which has nx in the n-th component for all n. Each Y_n is clearly in the radical of D; but Y is not for it is not even quasi-regular. Its quasi-inverse would have to be an element $Z = (\beta_1, \beta_2, \cdots, \beta_n, \cdots)$ with β_n the quasi-inverse of nx. But $\beta_n = -nx + n^2x^2 - n^3x^3 \cdots$ so that $|| \beta_n || = e^n - 1$. Hence Z is not in D. Thus the radical elements Y_n approach an element not in the radical; the radical of D is not closed.

To conclude this section we shall note some generalizations of theorems which are well known in the discrete case.

LEMMA 8. *If in a Q_r-ring the set $\{x + ax\}$, for fixed a and variable x, is dense, then a is* r. q. r.

Proof. We have $x + ax = -a + y$ with y arbitrarily small. Hence we may take y r. q. r., say $yoz = 0$. Then $aoxoz = yoz = 0$.

THEOREM 5. *A Q_r-ring with no proper closed right ideals is either a radical ring or a division ring.*

Proof. If A is not a radical ring, suppose that a is not r. q. r. By Lemma 8 the right ideal $\{x + ax\}$ cannot be dense in A. Hence $x + ax = 0$ for all x and $-a$ is a left unity element. Take a maximal right ideal M in A. If $M \neq 0$, M is both dense in A and closed by Theorem 2, a contradiction. Hence A has no proper right ideals at all. Such a ring is known to be a division ring.

We remark that the hypothesis that A is a Q_r-ring cannot be dropped in Theorem 5. An example is furnished by the ring of all rational numbers which can be written in the form ap^n with a an integer prime to p and n any integer, under the p-adic topology. This ring has no proper closed ideals but it is not a field.

THEOREM 6. *If a Q_r-ring has no proper closed ideals, its center is either a radical ring or a field.*

Proof. Suppose that the center contains an element a which is not q. r. It follows, as in the proof of Theorem 5, that $-a$ is a unity element 1. Now for any $b \neq 0$ in the center, bA is dense, so that $bx = 1 + y$ with $yoz = 0$, $bx(1 + z) = 1$, whence the center is a field.

If an element a in the center is q. r., then its (unique) quasi-inverse b is also in the center. For $xoa = aox$ for every x, and applying b left and right we obtain $box = xob$ whence b is in the center. It follows that if the center consists of q. r. elements, it is a radical ring.

4. Boundedness. Following Shafarevich [16] we shall say that a subset S of a topological ring A is *right bounded* if for any neighborhood U of 0 there exists a neighborhood V such that $V \cdot S \subset U$.[5] Left boundedness is similarly defined, and a set is bounded if it is both left and right bounded. We shall further say that A is locally (right) bounded if there exists in A an open set which is (right) bounded.

It is evident that a ring which has a neighborhood system consisting of (right) ideals is (right) bounded. This remark suggests the following example of a right bounded ring which is not left bounded: the ring of infinite matrices over a field having only a finite number of non-zero elements in each row, under the finite topology (a neighborhood of 0 consists of all matrices with first n rows 0). This ring is right bounded but not left bounded; in fact it is not even locally left bounded.

It is not true conversely that a bounded ring has ideal neighborhoods of 0; for a counter-example take any connected abelian group and make it a ring by defining all products to be 0. We can however prove the following partial converse.

LEMMA 9. *If a (right) bounded ring has a system of group neighborhoods of 0, then it has a system of (right) ideal neighborhoods of 0.*

Proof. Suppose that A is right bounded. Given a neighborhood U, let V be a group neighborhood with $V + V \subset U$, and W a group neighborhood with $W \subset V$, $WA \subset V$. Then $W + WA$ is an open right ideal contained in U. If A is bounded we use a similar argument for $W + WA + AW + WAW$.

[5] We use $A \cdot B$ to mean the set of all products ab, AB to mean the set of all sums of such products.

In case A is a metric ring, boundedness with respect to the metric implies boundedness as we have defined it, but not necessarily conversely: take the reals with all products equal to 0. However in a ring with a valuation the two notions coincide.

We list some further properties of boundedness which may be readily verified by the reader.

(1) Any discrete ring is bounded.
(2) Any subset of a bounded set is bounded.
(3) The closure of a bounded set is bounded.
(4) Any convergent sequence is bounded.
(5) The set-theoretic union of a finite number of bounded sets is bounded.
(6) If S and T are bounded, so are $S+T$ and $S \cdot T$.
(7) If $a_i \to 0$ and $\{b_i\}$ is bounded, then $a_i b_i \to 0$.
(8) The Cartesian sum of bounded rings is bounded.

The concept of boundedness derives much of its importance from the following result.

LEMMA 10. *Any compact set is bounded.*

Proof. Given a neighborhood U and any point x in the compact set K, we may find neighborhoods $V(x)$, $W(x)$ of x and 0 such that $V(x) \cdot W(x) \subset U$. A finite number of the V's cover K and if X is the intersection of the corresponding W's we have $KX \subset U$.

We shall now prove several theorems concerning bounded rings. It of course follows from Lemma 10 that these results extend at once to compact rings.

THEOREM 7. *The radical of a right bounded Q_r-ring is open.*

Proof. Let U be a neighborhood of 0 consisting of r. q. r. elements, and V a neighborhood with $V \cdot A \subset U$. Then for $x \in V$ and any $a \in A$, xa is r. q. r. Hence V is contained in the radical of A, which is therefore open.

COROLLARY 1. *A right bounded semi-simple Q_r-ring is discrete.*

COROLLARY 2. *A compact semi-simple Q_r-ring is finite.*

THEOREM 8.[6] *If C is the component of 0 in a right bounded locally compact ring A, then $CA = 0$.*

[6] This generalizes a result given by Otobe [14].

Proof. For any fixed element f in the character group of A, let $I(f)$ denote the set of all $a \in A$ with $f(aA) = 0$. We show that $I(f)$ is open. Choose neighborhoods U, V with $f(U) < 1/2$, $V \cdot A \subset U$. Then for $x \in V$ we have $nxA \in U$ for all integers n and hence $f(nxa) = nf(xa) < 1/2$ for all $a \in A$. It follows that $f(xa) = 0$, $x \in I(f)$, and $I(f)$ is open. Since $I(f)$ is a group, it is also closed. We then have $C \subset I(f)$, $f(CA) = 0$ for all f, whence $CA = 0$.

THEOREM 9. *A locally compact bounded ring with no proper closed ideals is discrete.*

Proof. The component of 0 in the given ring A is a closed ideal which must be either 0 or A. Hence A is either connected or totally disconnected. In the latter case it has group neighborhoods [13, 450], hence by Lemma 9 ideal neighborhoods of 0. This is possible only if A is discrete.

If A is connected, then by Theorem 8 we have $A^2 = 0$; in effect A is a connected locally compact abelian group with no proper closed subgroups. It follows [18, 110] that A is compact. But then the character group is discrete and has no proper subgroups at all; it must be finite which means that A is finite.

The theorem [9, Lemma 1] that a compact field is finite may now be generalized as follows.

COROLLARY. *A compact ring with no proper closed ideals is finite.*

Still another generalization is the following theorem.

THEOREM 10. *A right bounded division ring is discrete.*

Proof. In a division ring A we have $1 \in U \cdot A$ for any $U \neq 0$. Hence A can be right bounded only if it is discrete.

5. Compact semi-simple rings. We shall say that an element x is nilpotent if $x^n \to 0$. (In the present context there will be no confusion with purely algebraic nilpotence which of course corresponds to the discrete case.) A ring is a nil-ring if all its elements are nilpotent, and it is nilpotent if for any neighborhood U of 0 there exists N such that $R^n \subset U$ for $n > N$.

A nilpotent ideal is not necessarily in the radical: in the ring of integers under the p-adic topology the radical is 0, but the ideal (p) is nilpotent. However under suitable further assumptions the result is true.

THEOREM 11. *In a Q_r-ring, if a subsequence of $\{x^n\}$ approaches 0, then x is* r. q. r.

Proof. For some large n we have $(-x)^n$ r. q. r., say $(-x)^n oy = 0$. Set $z = [-x + x^2 - x^3 \cdots + (-x)^{n-1}]oy$. Then $xoz = (-x)^n oy = 0$.

COROLLARY. *In a Q_r-ring the radical contains all nil right ideals.*

THEOREM 12. *If $x^n \to 0$ in a complete ring with group neighborhoods of 0, then x is* q. r.

Proof. The series $-x + x^2 - x^3 + \cdots$ converges and, as in the proof of Lemma 3, its sum is a quasi-inverse of x.

COROLLARY. *In a complete ring with group neighborhoods of 0, the radical contains all nil right or left ideals.*

For compact rings it is possible to say a good deal more. We first prove that the radical is closed.

THEOREM 13. *In a compact ring the* r. q. r. *elements are closed.*

Proof. Suppose that x is not r. q. r., i. e., for all y, $xoy \neq 0$. We can find neighborhoods U_x, V_y of x and y such that $0 \notin U_x o V_y$. A finite number of the V_y's cover the ring. If U is the intersection of the corresponding U_x's, then U is a neighborhood of x which contains no r. q. r. elements.

From Theorem 4 we obtain the desired corollary.

COROLLARY. *The radical of a compact ring is closed.*

THEOREM 14. *The radical R of a compact totally disconnected ring A is nilpotent.*

Proof. Let B be an ideal neighborhood of 0. Since a homomorphism preserves quasi-regularity, the radical of $A - B$ contains all $r + B$ with $r \in R$. Also, since $A - B$ is finite its radical is nilpotent. Hence for large n we have $R^n \subset B$, as desired.

Because of the lack of group neighborhoods of 0, for a general compact ring we can only make the following weaker assertion.

THEOREM 15. *For any neighborhood U of 0 in a compact ring there exists N such that $R \cdot R \cdots \cdot R$ (n factors) $\subset U$ for $n > N$.*

Proof. Let C be the component of 0, and V a neighborhood with $V \cdot V \subset U$. The radical $R - C$ of $A - C$ is nilpotent, whence $R^m \subset V + C$ say for $m \geqq M$, and we need only take $N = 2M$.

Combining Theorems 12 and 15 we obtain the following result.

COROLLARY. *In a compact ring the radical is the union of all nil right and left ideals.*

For later use we now prove the following lemma.

LEMMA 11. *In a ring A with the descending chain condition on left ideals, the radical R contains the intersection of the maximal two-sided ideals.*

Proof. The ring $A - R$ is a semi-simple ring with descending chain condition, and so the intersection of its maximal ideals is 0. Suppose $a \notin R$; then there exists in $A - R$ a maximal ideal M not containing $a + R$. The inverse image in A of M is a maximal ideal not containing a.

We now prove the following conclusive structure theorem.

THEOREM 16. *A compact semi-simple ring is isomorphic and homeomorphic to a Cartesian direct sum of finite simple rings*

Proof. If C is the component of 0 in our ring A, we have $C^2 = 0$ by Theorem 8. Since A is semi-simple, $C = 0$ and A is totally disconnected. By Lemma 9, A has ideal neighborhoods of 0. For such an open ideal B, $A - B$ is compact and discrete and hence finite, and so contains an open maximal ideal whose inverse image in A is likewise an open maximal ideal.

Having thus proved the existence of open maximal ideals, we next show that the intersection of all of them is 0. Suppose on the contrary that $a \neq 0$ is in every open maximal ideal; the same will be true of ax for any x. Take an open ideal B. In $A - B$, the coset $ax + B$ will lie in every maximal ideal and so, by Lemma 11, will be nilpotent. We thus have $(ax)^n \in B$ for large n. By Theorem 12, a is in the radical of A, a contradiction.

Choose a fixed well-ordering $\{M_\lambda\}$ of the open maximal ideals. We select a subset as follows. Take $N_1 = M_1$, and, having chosen N_ρ for $\rho < \lambda$, take N_λ to be the first M not containing the intersection of a finite number of previously selected N's. When the process concludes we evidently have a meet-independent [2, 64] set $\{N_\lambda\}$ with intersection 0. It is immediate that we have a one-one representation of A as a subring of the direct sum of the rings $A - N_\lambda$. Moreover the mapping is continuous: a neighborhood of 0 in the direct sum consists of all elements having 0 at a finite number of places, and the intersection of the corresponding maximal ideals is an open set mapping into it.

For any finite number of N's we have [2, Th. 3.21] that $A - (N_{i_1} \cap \cdots \cap N_{i_k})$ is the direct sum of $A - N_{i_1}, \cdots, A - N_{i_k}$. Hence in the representation of A any finite combination of coordinates occurs, i. e.,

we have a dense subring of the direct sum. But the continuous image of a compact set is closed; hence we have the full direct sum. Since both spaces are compact, the mapping is also continuous in the reverse direction. Finally each $A - N_\lambda$ is certainly a finite ring with no proper ideals; it must be simple for if it were a zero ring A would not be semi-simple.

Remarks. 1. In the light of the revealed structure of A it is apparent that the N's are in fact the only open (or for that matter closed) maximal ideals in A. Thus the process of selecting the N's from the M's can be seen in retrospect to have been vacuous.

The following are fairly direct corollaries of Theorem 16.

2. A compact semi-simple ring has a unity element.

3. A compact semi-simple ring which has the ascending chain condition on closed ideals, or the descending condition on open ideals, is finite.

4. In a compact ring left and right quasi-inverses coincide (for this is true in any ring if it is true modulo the radical).

5. In a compact ring with unity element the radical is the intersection of the open maximal ideals.

6. A compact semi-simple ring satisfying the second axiom of countability is the direct sum of a countable number of finite simple rings.

7. A compact primitive ring (in the sense of Jacobson [11]) is finite.

6. Compact rings with radical. Since a finite simple ring is a matrix ring over a finite field, Theorem 16 assures us that a compact semi-simple ring has an ample supply of idempotents. We now require a device for transferring these idempotents to a ring with radical.

LEMMA 12. *Let A be a compact ring, B a nilpotent ideal in A, and H the homomorphism from A to $A - B$. Suppose that in $A - B$ we have a well-ordered set of idempotents $\{f_\rho\}$ such that $f_\alpha f_\beta = f_\beta f_\alpha = f_\alpha$ for $\alpha \leqq \beta$, and for λ a limit ordinal, $f_\lambda = \lim_{\rho<\lambda} f_\rho$. Then there exists in A a set of idempotents $\{F_\rho\}$ with $HF_\rho = f_\rho$, and $F_\alpha F_\beta = F_\beta F_\alpha = F_\alpha$ for $\alpha \leqq \beta$.*

Proof. Suppose that F_ρ has been found with the desired properties for $\rho < \lambda$.

Case I. λ not a limit ordinal. Let c be any element with $Hc = f_\lambda$. Define

$$c_1 = c - cF_{\lambda-1} - F_{\lambda-1}c + F_{\lambda-1}cF_{\lambda-1}.$$

Then

$$(4) \qquad Hc_1 = f_\lambda - f_{\lambda-1}, \qquad c_1F_\alpha = F_\alpha c_1 = 0 \qquad (\alpha \leqq \lambda - 1).$$

Define inductively $c_{n+1} = 3c_n^2 - 2c_n^3$. Then

$$(5) \qquad c_{n+1}^2 - c_{n+1} = 4(c_n^2 - c_n)^3 - 3(c_n^2 - c_n)^2.$$

We have $H(c_1^2 - c_1) = 0$, i. e., $c_1^2 - c_1 \in B$; by induction it follows from (5) that $c_n^2 - c_n \in B^{2^n}$. Let d be a limit point of $\{c_n\}$; then d is an idempotent and moreover, since each c_n shares properties (4) with c_1, so does d. On taking $F_\lambda = F_{\lambda-1} + d$ we successfully continue the induction.

Case II. λ a limit ordinal. Let F_λ be a limit point of F_ρ $(\rho < \lambda)$. By continuity $HF_\lambda = f_\lambda$ and $F_\alpha F_\lambda = F_\lambda F_\alpha = F_\alpha$ for $\alpha \leqq \lambda$, as desired.

We shall define a ring to be (*completely*) *primary* if it has a unity element, and modulo its radical is simple (a division ring). It is to be observed that for commutative rings the two notions coincide. We can prove the following structure theorem.[7]

THEOREM 17. *A commutative compact ring is the Cartesian direct sum of a radical ring and primary rings.*

Proof. Let A be the compact ring, R its radical, C the component of 0. We have that $C^2 = 0$ (Theorem 8) and the radical $R - C$ of $A - C$ is nilpotent (Theorem 14). By two successive applications of Lemma 12 we find in A an idempotent e mapping on the unity element of $A - R$. The Pierce decomposition relative to e expresses A as the direct sum of a radical ring and a ring with unity element. Our further discussion may therefore be confined to the latter, i. e., we assume that A has a unity element and is therefore totally disconnected.

The ring $A - R$ is a direct sum of fields, say with generating idempotents $\{e_\rho\}$. Define $f_\lambda = \sum_{\rho<\lambda} e_\rho$, apply Lemma 12, and finally define $E_\lambda = F_{\lambda+1} - F_\lambda$. The E's form a set of mutually orthogonal idempotents,

[7] For rings with ascending chain condition, an analogous theorem has been given by v. Dantzig [4, 213]. Jacobson's theorem [10, 442] that a compact ring is a direct sum of p-rings is both more and less general, since the ring need not be commutative while on the other hand p-rings are in general not primary.

and each AE_λ is a compact primary ring. The correspondence $a \to \{aE_\lambda\}$ maps A into the direct sum of the AE's. It is a one-one correspondence, since an element which annihilates each E_λ also annihilates each F_λ, and hence also the final F which is 1. The mapping is continuous, for if the directed set $\{a_\alpha\}$ converges to a, then $a_\alpha e$ converges to ae for any element e. Any finite combination of coordinates arises; hence we get the full direct sum. Finally since both spaces are compact, the mapping is continuous in the reverse direction.

It is well known that Theorem 17 is not valid for non-commutative rings, even in the discrete case. One must be content with a decomposition into left ideals, or into an additive direct sum of primary rings [5, 16-18]. Analogues of these two decompositions can be proved without difficulty for compact rings. It is also possible to prove the following theorem.[8]

THEOREM 18. *A compact primary ring is a matrix ring over a compact completely primary ring.*

We require the following lemma.[9]

LEMMA 13. *Let e be an idempotent in a compact ring A with radical R. Then eRe is the radical of eAe.*

Proof. In any ring we always have that eRe is contained in the radical of eAe. For given $r \in eRe$, we need only show that for any a, rae is r. q. r. in eAe. But if $(rae)ob = 0$, then $(rae)o(ebe) = 0$.

Conversely eRe is a nil ideal which by Theorem 15 is contained in the radical of eAe.

Proof of Theorem 18. Let A be the ring, R its radical; $A - R$ is a matrix ring over a finite field, say with matric units e_{ij} $(i, j = 1, \cdots, n)$. By Lemma 12, we have in A orthogonal idempotents $E_1, \cdots, E_n$ mapping on $e_{11}, \cdots, e_{nn}$.

We next show that there exist elements a, $b \in A$ such that $aE_1b = E_i$. The proof of this merely requires a re-examination of the construction used in Lemma 12. Let x, y be any elements mapping on e_{i1}, e_{1i} respectively. If necessary we can replace x by E_ix, y by yE_i, and so we may assume $E_ix = x$,

[8] Our proof is actually valid for any complete topological ring whose radical is nilpotent, and which modulo the radical is a simple ring with minimum condition.

[9] Lemma 13 will hold under any hypothesis which assures us that the radicals of A and eAe are the union of all nil-ideals, for example A may be a normed ring.

$yE_i = y$. We take xE_1y to be the c_1 of Lemma 12. Then $c_2 = 3c_1^2 - 2c_1^3 = xE_1y_2$ where $y_2 = (3 - 2yxE_1)yxE_1y$, and in general $c_n = xE_1y_n$. Now c_n approaches an idempotent E, and if b is a limit point of $\{y_n\}$, we have $E = xE_1b$. Evidently $bE_i = b$ and so $EE_i = E_iE = E$. Also E and E_i both map on e_{i1}. Hence $E - E_i$ is an idempotent contained in R, whence $E = E_i$.

Now with $E_i = aE_1b$, define $E_{1i} = E_1b$, $E_{i1} = aE_1$, $E_{ij} = E_{i1}E_{1j}$. Then $E_{i1}E_{1i} = E_{ii} = E_i$. Also $E_{1i}E_{i1}$ is an idempotent mapping on e_{11} and is invariant under multiplication by E_1. Hence $E_{1i}E_{i1} = E_1$. It then follows that $E_{ij}E_{kl} = \delta_{jk}E_{il}$.

Consider the right ideal E_1A. The mapping $a \to ax$ gives for every $x \in A$ an endomorphism of E_1A. The correspondence is one-one, for if $E_1Ax = 0$, then $E_ix = 0$ and $x = 0$, since $1 = E_1 + \cdots + E_n$. Let T be an endomorphism of E_1A commuting with all these right multiplications, i. e., $(E_1a)Tx = (E_1ax)T$ for all a and x. Then $E_1T = (E_1T)E_1 = E_1bE_1$, say. Also $(E_1a)T = (E_aT)a = (E_1bE_1)(E_1a)$; in short, T is a left multiplication by an element in E_1AE_1.

We now regard E_1A as a left vector space over E_1AE_1. It is easy to see that $E_1, E_{12}, \cdots, E_{1n}$ form a basis for this vector space; for

$$E_1a = E_1a \cdot \Sigma E_i = \Sigma(E_1aE_{i1}E_1)E_{1i},$$

while $\Sigma a_iE_{1i} = 0$ implies $a_iE_1 = 0$ on right multiplication by E_{i1}.

We have represented A as a set of endomorphisms on this vector space. We get all endomorphisms, since E_{ij} sends E_{1i} into E_{1j} and the other basis elements into zero. By Lemma 13, E_1AE_1 is completely primary and since it is closed it is also compact.

It remains finally to identify the topology of A with that of the matrix ring. A neighborhood U of 0 in the latter consists of matrices with elements in a neighborhood V of 0 in E_1AE_1. Take a neighborhood W in A such that $E_{1i}xE_{j1} \in V$ for all i, j and all $x \in W$. Then the matrix representing x will lie in U. This proves continuity of the mapping one way; the reverse direction follows since both spaces are compact.

We can also establish the following result.

THEOREM 19. *A compact ring with no divisors of zero is either a radical ring or completely primary.*

Proof. In a ring with no divisors of zero there is at most one idempotent—the unity element. It follows from Lemma 12 that the ring modulo its radical is either zero or a division ring.

A commutative primary ring has a radical consisting precisely of all non-units, and will thus be a local ring if the ascending chain condition is fulfilled. We shall now give an alternative sufficient condition for this purpose.

THEOREM 20. *Let A be a compact commutative primary ring with radical R. If R^2 is open, then A is a local ring and has the latter's natural topology.*

Proof. The ring $R - R^2$ is compact and discrete and hence finite. Let $a_1, \cdots, a_n$ be any representatives in R of these cosets, and let $B = (a_1, \cdots, a_n)$. Clearly $B \subseteq R$; we assert that $R = B$. Any element in R is expressible as a sum of elements in B and R^2, and more generally one in R^i as a sum of elements in $R^{i-1}B$ and R^{i+1}. Since for any neighborhood U, $R^n \subset U$ for large n, it follows that $r \in R$ can be written $r = \sum_{i=1}^{\infty} r_i$ with r_i in $R^{i-1}B$. Write $r_i = \Sigma c_{ij}a_j$ $(c_{ij} \in R^{i-1})$ and we see that $\Sigma_i c_{ij}$ converges say to c_j', $r = \Sigma c_j'a_j \in B$.

We have thus shown that R has a finite basis. It now follows from a result due to Cohen [3, Th. 3] that A is a local ring. Since the residue class field is finite, A is compact in the local ring topology (this topology is defined by taking $\{R^n\}$ to be a system of neighborhoods of 0). The mapping

$$A\,(\text{local ring topology}) \rightarrow A\,(\text{original topology})$$

is continuous since any neighborhood of 0 contains some R^n. Since A is compact in both topologies, the mapping is a homeomorphism.

We shall now give an example of a compact integral domain which is not a local ring. Let A_m $(m = 1, 2, \cdots)$ denote the ring of all formal power series in $x_1, \cdots, x_m$ with coefficients in some finite field. For $m < n$ there is a natural homeomorphism θ_{mn} of A_n into A_m obtained by simply suppressing the variables $x_{m+1}, \cdots, x_n$. Relative to the projections θ_{mn}, $\{A_m\}$ forms an inverse system [6, § 20]. Each A_m is a compact integral domain and so the limit ring is again a compact integral domain. However A is not a local ring since, for example, $(x_1) \subset (x_1, x_2) \subset (x_1, x_2, x_3) \subset \cdots$ is an ascending chain of ideals. It may also be observed that R^2 is not open since any neighborhood of 0 contains $x_n \notin R^2$ for large n.

We shall conclude this section with some remarks on locally compact rings.

LEMMA 14. *A locally compact ring without divisors of zero is either connected or totally disconnected.*

Proof. Let C be the component of 0. C is a locally compact connected ring without divisors of zero, hence [9] C is either 0 or a division algebra over the reals. In the latter event let 1 be the unity element of C. Then for any $a \in A$, $1 \cdot a = c \in C$. But then $1(a - c) = 0$, $a = c$.

LEMMA 15. *Let x be an element in a ring with the second axiom of countability and without divisors of zero such that the closure of $\{x^n\}$ is compact. Then either $x^n \to 0$ or $\{x^n\}$ has a subsequence approaching* 1.

Proof. Suppose $x^{n(i)} \to a \neq 0$. A subsequence of $\{x^{n(i+1)-n(i)}\}$ approaches y and $ay = a$. Then for any z, $a(yz - z) = 0$. If $a \neq 0$, $yz = z$ and similarly $zy = z$, $y = 1$.

LEMMA 16. *In a locally compact ring with the second axiom of countability and without divisors of zero there exists a neighborhood of* 0 *consisting of nilpotent elements.*

Proof. Let U be a compact neighborhood of 0; if A has a unity element 1 we may assume $1 \notin U$. Let V be a neighborhood of 0 with $V \subset U$, $VU \subset U$. Then for $x \in V$, $x^n \in U$. By Lemma 15, $x^n \to 0$.

LEMMA 17. *For a locally compact totally disconnected ring A with the second axiom of countability the following three statements are equivalent:*

(1) *A is a Q-ring,*

(2) *A is a Q_r-ring,*

(3) *A has a neighborhood of* 0 *consisting of nilpotent elements.*

Proof. That (1) implies (2) is obvious, and that (3) implies (1) follows from Theorem 14. To prove that (2) implies (3) we proceed as in Lemma 15. When we reach $ay = a$ we may assume that $-y$ is r. q. r., say $-yoz = 0$. Left-multiplying by a, we obtain $ay = 0$, $a = 0$.

THEOREM 21. *A locally compact ring with the second axiom of countability and without divisors of zero is a Q-ring.*

Proof. If A is connected it is a division ring, and any division ring is a Q-ring since -1 is the only element which is not q. r. If A is totally disconnected, the theorem follows from Lemmas 16 and 17.

For the case of a division ring the following result was proved by Otobe [15].[10]

THEOREM 22. *In a locally compact ring with the second axiom of countability and without divisors of zero the quasi-inverse is continuous.*

Proof. This is true for the reals, complexes, or quaternions, so we assume that A is totally disconnected. By Lemma 4 and Theorem 21 it suffices to prove continuity at 0. By taking x in the neighborhood V of the proof of Lemma 16, we have that its quasi-inverse $-x+x^2-x^3\cdots$ is in U.

PART II. Rings of Functions.

7. Functions on totally disconnected spaces. Let $C(X, A)$ denote the ring of continuous functions from a topological space X to a topological ring A. We shall consider three ways of topologizing $C(X, A)$, listed in order of increasing strength:

p: pointwise convergence,

k: uniform convergence on compact subsets of X,

u: uniform convergence on all of X.

Of course p and k coincide if S is discrete, and k and u coincide if X is compact. When desirable we shall indicate the choice of topology by the notation $C(X, A, p)$, etc.

Some elementary facts about $C(X, A)$ are assembled in the following theorem.

THEOREM 23. *Suppose that the quasi-inverse in A is continuous where defined. $C(X, A)$ can be a Q-ring only if A is a Q-ring, and if the latter is, so is $C(X, A, u)$ but not in general $C(X, A, p$ or $k)$. The radical of $C(X, A)$ is $C(X, R)$ where R is the radical of A; it is closed under the p-topology (or any stronger topology) if and only if R is closed. $C(X, A, u)$ is complete if and only if A is. If X is locally compact, $C(X, A, k)$ is also complete.*

---

[10] The elements with quasi-inverses form a group under o which is complete in a suitable metric. Hence Theorem 22 can also be derived from a result due to Montgomery: "Continuity in topological groups," *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 879-882. (I am indebted to the referee for this remark.)

Proof. Let $f \in C(X, A)$ be such that $f(x)$ is q. r. for all x. There is a unique function g with $f(x)og(x) = 0$ for all x. By the continuity of the quasi-inverse, g is continuous. It follows that if A is a Q-ring, so also is $C(X, A, u)$. Under the p- or k-topology, neighborhoods of 0 will in general contain functions which assume values not q. r.

We leave to the reader the remainder of the proof of Theorem 23.

The connection between the structure of $C(X, A)$ and that of X and A has been widely studied for the case where A is the reals or complexes [8], and for the case where X is totally disconnected and A is the field of integers mod 2 [17], or more generally a discrete division ring [12]. We shall proceed to study the case where X is totally disconnected under less restrictive assumptions on A; however we shall assume that A has a unity element.

We shall use the following terminology suggested by E. Hewitt: an ideal I in $C(X, A)$ is *free* if for every $x \in X$ there exists $f \in I$ with $f(x) = 1$; otherwise I is *fixed.*

THEOREM 24. *If X is totally disconnected and compact and A is a Q-ring with unit and inverse continuous where defined, then every proper right ideal in $C(X, A)$ is fixed. A maximal (right or two-sided) ideal M has the following form: for some $x \in X$ and maximal right or two-sided ideal N in A, M consists of all f with $f(x) \in N$.*

Proof. Suppose that I is a free right ideal, and let $f \in I$ be a function with $f(x) = 1$. In a suitable neighborhood U_x of x, which we may take to be open and closed, we have that f^{-1} exists. Define a function g as follows: $g = f^{-1}$ in U_x and is 0 otherwise. The function g is continuous and so $h = fg \in I$; h is the characteristic function of U_x. A finite number of the U's cover X; by taking intersections and reselecting the corresponding h's, we may suppose that we have disjoint open and closed sets $U_1, \cdots, U_n$ covering X, with characteristic functions $h_1, \cdots, h_n$ in I. Then $h_1 + \cdots + h_n = 1$ is also in I and $I = C(X, A)$. The final statement of the theorem follows from the fact that quite generally any fixed maximal ideal has the designated form.

THEOREM 25. *Suppose that $C(X, A)$ is isomorphic to $C(X_1, A_1)$ where X, X_1 are totally disconnected and compact and A, A_1 are Q-rings with unit elements. Then for any maximal ideal N in A there exists a maximal ideal N_1 in A_1 such that $A - N$ and $A_1 - N_1$ are isomorphic. If A, A_1 are simple, then they are isomorphic and X, X_1 are homeomorphic.*

Proof. Take a point $x \in X$ and construct the maximal ideal M of functions f with $f(x) \in N$. Let M_1 be the corresponding maximal ideal in $C(X_1, A_1)$. By Theorem 24 there exists in A_1 a maximal ideal N_1 such that $A - N \cong C(X, A) - M \cong C(X_1, A_1) - M_1 \cong A_1 - N_1$.

If A, A_1 are simple, then $N = N_1 = 0$, so that $A \cong A_1$. Moreover there is a one-one correspondence between points of X and maximal ideals in A. If we introduce the Stone topology into the latter, we verify by well known methods that the correspondence is a homeomorphism (see for example [8]).

Theorem 25 can be interpreted as saying that when X is compact and totally disconnected and A is simple, the algebraic properties of $C(X, A)$ characterize X completely. This is no longer true if the hypothesis of compactness is dropped. However we shall now show that the algebraic and topological properties of $C(X, A, k)$ together do characterize X. At the same time we are able to weaken slightly the hypothesis on A.

THEOREM 26. *Let X be totally disconnected. Any closed (right) ideal I in $C(X, A, k)$ has the following form: for every $x \in X$ a closed (right) ideal $J(x)$ in A is prescribed and I consists of all functions f with $f(x) \in J(x)$ for all x.*

Proof. Define $K(x)$ to be the set of values in A assumed at x by functions in I and let $J(x)$ be the closure of $K(x)$. Let f be a function with $f(x) \in J(x)$ for every x; we shall show that $f \in I$. Take a neighborhood of 0 in $C(X, A, k)$, consisting say of all functions mapping a compact set K into a neighborhood U of 0 in A. For any $x \in K$ there is a function $g \in I$ with $g(x) - f(x) \in U$, and this will extend to an open and closed neighborhood U_x of x. We may assume that g vanishes outside U_x. A finite number of the U's cover K; we may assume them non-overlapping. The sum $h \in I$ of the corresponding g's then has the property that $h(x) - f(x) \in U$ for all x in K. Hence I contains functions arbitrarily close to f, and since it is closed it must contain f.

COROLLARY 1. *If in addition A has no proper closed ideals, then any closed ideal in $C(X, A, k)$ consists of all functions vanishing on some (necessarily closed) set of X. In particular a closed maximal ideal consists of all functions vanishing at a point.*

COROLLARY 2. *Let X, X_1 be totally disconnected spaces and A, A_1 rings with no proper closed ideals such that $C(X, A, k)$ is isomorphic and homeo-*

morphic to $C(X_1, A_1, k)$. Then A, A_1 are isomorphic and homeomorphic, and X, X_1 are homeomorphic.

Suppose that the maximal right ideals in A are closed (this is true by Theorem 2 if A is a Q_r-ring); their intersection is the radical R of A. It follows readily from Theorems 23 and 26 that the intersection of the closed maximal right ideals of $C(X, A, k)$ already gives the radical of that ring. This is not true for all rings, since in **3** we gave an example of a ring with non-closed radical.

It is natural to ask: what about the free maximal ideals? For locally compact spaces we can give the following result.

THEOREM 27. *Let X be a locally compact totally disconnected space and A a Q-ring with unit; and let F consist of all $f \in C(X, A)$ such that the closure of the set of points x for which $f(x) \notin R$ (the radical of A) is compact. Then a maximal right ideal in $C(X, A)$ is free if and only if it contains F.*

Proof. The ideal F is itself free; indeed it is easy to see that it is dense in $C(X, A, k)$. Thus any ideal containing F is free. Conversely suppose that M is a free maximal right ideal and that $f \in F$; we show that $f \in M$. The set of points x where $f(x) \notin R$ can be embedded (since X is locally compact) in an open compact set K. The function g, defined by $g(x) = 0$ in K and $g(x) = f(x)$ outside K, lies in the radical of $C(X, A)$; hence $g \in M$. Thus to prove $f \in M$ it will suffice to prove that $f - g \in M$, and for this it will in turn suffice to prove $k \in M$, where k is the characteristic function of K; this last fact follows from Theorem 24.

Theorem 27 shows that the intersection of the free maximal right ideals contains F. Whether or not it is equal to F is an open question. A partial affirmative answer is afforded by the following result.

THEOREM 28. *With the same hypothesis as in Theorem 27, suppose further that either X or A is discrete. Then F is the intersection of the free maximal right ideals of $C(X, A)$.*

Proof. Suppose that on the contrary there exists a function f not in F but in every free maximal right ideal. For any $a \in A$ we shall denote by $f^{-1}(a)$ the set of $x \in X$ with $f(x) = a$. By virtue of our assumption that either X or A is discrete, $f^{-1}(a)$ is open and closed. We define two further functions g and h, distinguishing two cases.

I. $a \in R$, the radical of A. Set $g(x) = 0$, $h(x) = 1$ for $x \in f^{-1}(a)$.

II. $a \notin R$. Then there exists a maximal right ideal N in A which with a generates A, i. e., there exist elements b, c with $ab + c = 1$, and such that c does not have a right inverse. We set $g(x) = b$, $h(x) = c$ for $x \in f^{-1}(a)$.

The method of definition assures us that g and h are continuous functions and that $fg + h = 1$.

We show next that h is contained in some free maximal right ideal. For this, by Theorem 27, it will suffice to prove that h and F generate a proper ideal. Suppose the contrary: then there exist functions p, q, with $q \in F$, such that $hp + q = 1$. Let P denote the set of points where $1 - q$ fails to have an inverse. By the definition of F, the closure of P is compact. Hence the set P' of points where h fails to have a right inverse has compact closure. By the definition of h, P' includes all x such that $f(x) \notin R$. Hence the points x where $f(x) \notin R$ have compact closure, and $f \in F$, a contradiction.

It must then be the case that f and h together generate a proper right ideal, an absurdity in view of $fg + h = 1$.

8. Prime ideals. Let P be a prime ideal in $C(X, A)$. It is assumed that X is totally disconnected. For every $x \in X$, the values assumed by functions of P at x form an ideal $I(x)$. We assert that, with the exception of at most one point, we must have $I(x) = A$. For suppose that at the point x and y the values a and b respectively are excluded. We surround x by an open and closed set U excluding y, and form functions f, g equal to a, b on U and the complement of U respectively, and zero otherwise. Then $fg = 0$ so that either f or g is in P, a contradiction. In case A is simple, this result may be restated as follows: a prime ideal can have at most one zero. If X is compact, P has precisely one zero, as follows from Theorem 24.

In our further investigation we make the following three assumptions on A:

(1) A has no divisors of 0,

(2) division where possible is continuous,

(3) for any $a \in A$ there exists a neighborhood of 0 consisting of right multiples of a.

(Examples of rings satisfying these postulates: any division ring, any discrete ring with no divisors of zero, the ring of p-adic integers, etc.)

It can be proved from these axioms that the closure Q' of a prime ideal

Q in A is prime. For suppose $ab \in Q'$ with neither a nor b in Q'. A sufficiently small neighborhood of ab consists of right multiples of a; thus there exist elements $ac \in Q$ arbitrarily near to ab. Since $a \notin Q$, we have $c \in Q$. By continuity of division, any neighborhood of b contains an element of Q. Hence $b \in Q'$, a contradiction.

Let P again be a prime ideal in $C(X, A)$, and suppose that at the point x it takes on the values of a proper ideal I. We assert that I is a prime ideal. Suppose $f(x) = ab \in I$. In a small open and closed neighborhood U, f will continue to be a right multiple of a, and the quotient by our assumption is continuous. Let $g = a$ on U and 1 on the complement; $h = a^{-1}f$ on U and f on the complement. Then $gh \in P$, so that either g or h is in P, whence either a or b is in I.

By Theorem 26 the closure of P in $C(X, A, k)$ consists of all functions with $f(x)$ in the closure of I. We may summarize these results in the following theorem.

THEOREM 29. *Suppose that X is totally disconnected and that A satisfies the three assumptions above. Then the closure of a prime ideal in A or $C(X, A, k)$ is prime. Further the following statement is true in $C(X, A, k)$ if and only if it is true in A: a closed prime ideal is either maximal or the whole ring.*

With more drastic assumptions it is possible to prove that prime ideals in $C(X, A)$ are maximal.

THEOREM 30. *Suppose that A is a division ring and that either X or A is discrete. Then the prime ideals in $C(X, A)$ are maximal.*

Proof. Let P be a prime ideal and suppose that (P, f) is proper, $f \notin P$. Let $Z \subset X$ denote the set of zeros of the function f; under either hypothesis Z is open and closed. Let g and h be the characteristic functions for Z and the complement of Z; since $gh = 0$ we have either g or h in P. In the first case $g + f \in P$ is a function with no zeros, which is impossible. In the second case $f = fh$ is already in P.

We shall now give an example to show that the hypotheses of Theorem 30 cannot be light-heartedly removed.[11] Suppose that A is a field with a valuation and that all non-negative real numbers are assumed as values (actually a dense

[11] This example is based on a similar one constructed by E. Hewitt for the case where A is the reals and X the unit interval.

set would suffice). Let us take X to be compact and totally disconnected, and let $X = U_1, U_2, \cdots$ be a sequence of compact open sets shrinking to the point x. Take any function f satisfying $|f| = 1/n$ on $U_n - U_{n+1}$, and another function g satisfying $|g| = e^{-n}$ on $U_n - U_{n+1}$. Then the ideal generated by g does not contain any power of f. For suppose $f^p = gh$, with $|h| \leq M$. Then on $U_n - U_{n+1}$ we have $n^{-p} \leq Me^{-n}$ which cannot be true for all n.

By Zorn's lemma, expand the ideal (g) to one which is maximal with respect to the exclusion of the powers of f. The result is a prime ideal, but not a maximal one since it lacks some of the functions vanishing at x.

9. Primitive ideals. In a ring with unit a primitive ideal [11] is the largest two-sided ideal contained in a maximal right ideal. From Theorem 24 the following result follows at once.

THEOREM 31. *Let X be compact and totally disconnected, and A a Q-ring. A primitive ideal in $C(X, A)$ consists of all functions f with $f(x) \in P$, where x is a fixed point of X and P is a primitive ideal in A.*

In particular we see that if all primitive ideals in A are maximal, the same will be true in $C(X, A)$.

Jacobson [12] has raised the following question. Let X be compact and totally disconnected, A a discrete division ring, and let C^* be a subring of $C(X, A)$ containing the constant functions and for any $x, y \in X$ a function with $f(x) \neq f(y)$. Are the primitive ideals in C^* maximal? We shall show that the answer is affirmative by proving that $C^* = C(X, A)$. Actually we can prove the following more general result (it is to be noted that in the above case $C(X, A, k)$ is discrete).[12]

THEOREM 32. *Let X be totally disconnected and A a ring having a system of ideal neighborhoods of 0. Let C^* be a subring of $C(X, A)$ containing the constant functions; and for any $x, y \in X$ a function f with $f(x) = 1$, $f(y) = 0$. Then C^* is dense in $C(X, A, k)$.*

Proof. Since we can work within a neighborhood of 0 in $C(X, A, k)$, it suffices to treat the case where X is compact. Let U be an ideal neighborhood of 0 in A, and let K be an open compact set in X, L its complement. We shall show that C^* contains a function f with $f(K) \subset 1 + U$, $f(L) \subset U$.

[12] For the case where A is a finite field, Theorem 32 was communicated to me by R. Arens.

This will complete the proof, for it is readily seen that any function in $C(X, A, k)$ can be approximated within U by a finite linear combination of such functions.

Take a fixed point $y \in L$. For any $x \in K$ there exists a function g with $g(x) = 0$, $g(y) = 1$. In a neighborhood V_x of x we have $g(V_x) \subset U$. A finite number of the V's cover K, and the product of the corresponding g's gives us a function h with $h(K) \subset U$, $h(y) = 1$. The function $1 - h$ takes values within U in a neighborhood W_y of y. A finite number of W's cover L, and the product of the corresponding $(1 - h)$'s gives us the desired function f.

The space of primitive ideals of a ring, topologized after Stone, has been called the *structure space* by Jacobson. In connection with this concept, we can prove the following result.

THEOREM 33. *If X is totally disconnected and compact, the structure space of $C(X, A)$ is the Cartesian product of X and the structure space of A.*

Proof. Theorem 31, together with its obvious converse, shows that there is a one-one correspondence between the two spaces in question; moreover each of them is compact. In the direct product take the set D of all pairs (b, p), $b \in B$, $p \in P$, where B is closed in X and P is a closed set of primitive ideals in A, say with intersection I. It will suffice to show that E, the corresponding set of primitive ideals in $C(X, A)$, is closed. Let J be the intersection of the ideals in E; clearly J consists of all f with $f(x) \in I$ for $x \in B$. Moreover for any $y \notin B$, J contains a function with $f(y) = 1$: we may surround B by an open and closed set U excluding y, and take $f = 0$ on U, $f = 1$ on the complement of U. It follows from Theorem 31 that any primitive ideal containing J must consist of all f with $f(z) \in Q$ for a suitable point z in B and a primitive ideal Q in A. Finally we observe that Q must contain I; since P is closed we have $Q \in P$, and it follows that E is closed.

10. Functions with values in certain division rings. Let X be a compact T_0-space and A a division ring. It appears to be an open question whether proper ideals in $C(X, A)$ are necessarily fixed. We have seen (Theorem 24) that the answer is affirmative when X is totally disconnected. Another case where the answer is known to be affirmative is that where A is the field of reals or complexes. This latter case can be generalized as follows.

Suppose that A is a division ring admitting a continuous function $x \to x^*$ such that $xx^* + yy^* = 0$ implies $x = y = 0$. From the case $x = y$ it is evident that such a function cannot exist in a ring of characteristic 2.

However one exists in every ring of characteristic different from 2 that I have examined.

Examples. (1) Discrete. Take $0^* = 0$, $x^* = x^{-1}$ for $x \neq 0$.

(2) Formally real. Take $x^* = x$.

(3) Complete in an archimedean valuation. Then A is the reals, complexes or quaternions, and we take x^* to be the conjugate of x.

(4) Non-archimedean valuation. A possible choice is the following. For each value m take a fixed element z_m of that value. Define $0^* = 0$, $x^* = z_m{}^2 x^{-1}$, if $x \neq 0$ has value m.

(5) The rational numbers under the 6-adic topology. Take $0^* = 0$, and if $r = 2^m 3^n s t^{-1}$ (s, t prime to 6), $r^* = 2^{2m} 3^{2n} r^{-1}$.

In the following we shall write f^* for the function defined by $f^*(x) = [f(x)]^*$, and $Z(f)$ for the set of zeros of f. It is assumed throughout that X is a T_0-space and that A possesses a *-function.

LEMMA 18. *For any functions $f_1, \cdots, f_n$ in $C(X, A)$ there exist functions $g_1, \cdots, g_n$ such that $Z(f_1 g_1 + \cdots f_n g_n) = Z(f_1) \cap \cdots \cap Z(f_n)$.*

Proof. By induction we may assume that we have $h_1, \cdots, h_{n-1}$ with $Z(k) = Z(f_1 h_1 + \cdots + f_{n-1} h_{n-1}) = Z(f_1) \cap \cdots \cap Z(f_{n-1})$. We then take $kk^* + f_n f_n{}^*$.

LEMMA 19. *If I is a free proper right ideal in $C(X, A)$, then for every $f \in I$, $Z(f)$ is non-compact.*

Proof. For any functions $f_1, \cdots, f_n$ in I, $Z(f_1) \cap \cdots \cap Z(f_n)$ must be non-void by Lemma 18. Suppose that for some $f \in I$, $Z(f)$ is compact. Then the intersection of all $Z(g) \cap Z(f)$, as g ranges over I, is non-void, and I is fixed.

COROLLARY. *If X is compact, every proper ideal in $C(X, A)$ is fixed.*

LEMMA 20. *If I is a free right ideal in $C(X, A)$ and $K \subset X$ is compact, then there exists in I a function not vanishing in K.*

Proof. For any x in K we have a function $f \in I$ and a neighborhood U_x such that $Z(f) \cap U_x$ is void. A finite number of the U's cover K, and by Lemma 18 we may build from the f's the desired function.

COROLLARY. *A closed proper right ideal in $C(X, A, k)$ is fixed.*

LEMMA 21. *If M is a maximal right ideal in $C(X, A)$, $f \in M$, and $Z(g) \supseteq Z(f)$, then $g \in M$.*

Proof. If not we have $gh + m = 1$ for a suitable $m \in M$. But then $Z(g) \cap Z(m)$, and *a fortiori* $Z(f) \cap Z(m)$, is void; by Lemma 18, $1 \in M$.

Lemma 21 carries with it the evident corollary that the maximal right ideals in $C(X, A)$ are two-sided and the primitive ideals maximal. However in this connection it may be of interest to show that not all right ideals are necessarily two-sided.

Example. Let X be the unit interval, A the quaternions. Consider the principal right ideal generated by $f = x + ixy$, where we have written y for $\sin(1/x)$. If this ideal is two-sided we must have $jf = fg$ for some continuous g. But for $x \neq 0$ we have

$$g = f^{-1}jf = [(1 - y^2)j - 2yk]/(1 + y^2),$$

and this cannot be extended to be continuous at $x = 0$.

Let M be a maximal ideal in $C = C(X, A)$. If M is fixed, $C - M$ is isomorphic to A. If M is free, $C - M$ is still a division ring containing A — for the constant functions must be sent into separate residue classes. The *-function can be extended to $C - M$, granted that $0^* = 0$; for then by Lemma 21, $f \in M$ implies $f^* \in M$, so that the *-function is uniquely definable on residue classes, and $ff^* + gg^* \in M$ implies $f, g \in M$, so that the fundamental property persists.

We shall now give an explicit instance where $C - M$ is a proper extension of A. Let X be the same abstract space as A, and let f be the function which maps X identically on A. It is impossible that $f - c \in M$ for c a constant function, for $f - c$ has only one zero and by Lemma 19 this would be a zero of all of M, making M a fixed ideal. Thus f does not map into A, and $C - M$ must be a proper extension of A.

11. Functions vanishing at a point. Let $C(X, A, x)$ be the subring of $C(X, A)$ consisting of all functions vanishing at the point $x \in X$. If X is compact we may alternatively describe the functions as vanishing "at infinity" on the locally compact space $X - x$. Rings of this kind occur in the theory of normed rings and Boolean rings without unit.

THEOREM 34. *If X is a compact Hausdorff space and either* (1) *A is the field of reals or complexes, or* (2) *X is totally disconnected and A is a finite field or a field of characteristic zero, then a maximal ideal M in $C(X, A, x)$ is fixed.*[13]

[13] We mean here that all the functions of M vanish at some point other than x.

Proof. First we prove that M is a prime ideal. The alternative is for $C - M = C(X, A, x) - M$ to be a zero ring multiplicatively and a group of p elements additively. But then if A has characteristic zero, any $f \in C$ is of the form $f = pg$, so $f \in M$ and $M = C$. If A is a finite field we have $f^{p^n} = f$ and again $M = C$.[14]

Suppose that M is a free ideal. Then for any $y \in X$, $y \neq x$, there is a function in M not vanishing at y. Following the method of Theorem 24 or Lemma 18 as the case may be, we can show that for any neighborhood of x, M contains a function not vanishing in its complement.

Next we show that M contains any function f vanishing in a neighborhood U of x. Let $g \in M$ be a function not vanishing in the complement V of U. Define $h = g^{-1}f$ in V, $h = 0$ in U. Then h is readily seen to be continuous, and $f = gh$ is in M.

Now suppose h is a function not contained in M. Since M is prime, also $h^2 \notin M$. Hence $(h^2, M) = C$, and $h = kh^2 + m$ for suitable $k \in C$, $m \in M$. Take a neighborhood U of x such that $hk \neq 1$ in the closure U' of U. Then $m/h = 1 - kh$ is a continuous function different from zero in U', and so is its reciprocal. This function h/m may be extended to a continuous function r defined everywhere in X; in case (1) this follows from Tietze's Theorem, in case (2) we may take U to be open and closed and set $r = 0$ outside U. The function hr is in C since it vanishes at x, and so hrm is in M. But $hrm - h^2$ vanishes in U and, by our above result, it is also in M. Hence $h^2 \in M$, a contradiction.

12. The ring of bounded functions. Let $C'(X, A)$ denote the ring of continuous bounded functions from X to A — boundedness being defined as in **4**. The following result is easy to verify.

THEOREM 35. *$C'(X, A, u)$ is a Q-ring if A is a locally bounded Q-ring; if further A is complete, so also is $C'(X, A, u)$.*

It is to be observed that $C'(X, A, k)$ is not in general complete. In fact, if A is the reals and X is normal, then $C'(X, A, k)$ is dense in $C(X, A, u)$.

Suppose that A is a division ring. The spaces of maximal ideals in $C(X, A)$ and $C'(X, A)$ both provide compactifications of X. Gelfand and Kolmogoroff [8] have shown for the case where A is the reals that these compactifications are homeomorphic. Their technique does not seem to be applicable in more general cases. By a different method we are able to prove the following theorem.

---

[14] The hypotheses on A in case (2) are only needed at this point to prove that M is prime. Still another adequate hypothesis is that A has a valuation.

THEOREM 36. *Let X be discrete and A a locally bounded division ring with the property that for any neighborhood U of 0 the set of inverses of the complement of U is bounded.*[15] *Then the spaces of maximal ideals in $C(X, A)$ and $C'(X, A)$ are homeomorphic.*

Proof. Let M be a maximal ideal in C. $N = M \cap C'$, being the contraction of a prime ideal, is prime, but in general is not a maximal ideal in C'. However it expands uniquely to a maximal ideal M'. To prove this we need only show that (N, f, g) is a proper ideal whenever (N, f) and (N, g) are both proper. Suppose on the contrary that $n + f + g = 1$, $n \in N$. Let U be a neighborhood of 0 in A such that $(1 + U)^{-1}$ is bounded (such a neighborhood exists because of local boundedness). Let Y be the intersection of $Z(n)$—the set of zeros of n—and the set where f takes values in U. Let p be the function equal to 0 on Y and 1 elsewhere, q the function equal to 0 on $Z(n) - Y$ and 1 elsewhere. On the set where p vanishes, g takes values in $1 + U$ and has a bounded inverse. Hence $(g, p) = C'$. On the set where q vanishes, f takes values in the complement of U and by our assumption has a bounded inverse. Hence $(f, q) = C'$. But $pq \in N$ so that either p or q must be in N. In either case we have a contradiction.

Conversely let us start with a maximal ideal M' in C'. For any f in C, define f' to be the function with the same zeros as f, but equal to 1 elsewhere, i. e., f' is the characteristic function of the complement of $Z(f)$. We define M as follows: $f \in M$ if and only if $f' \in M'$. It is readily verified that M is a proper ideal in C. Suppose that M is not maximal, so that (M, g) is proper, $g \notin M$. If g'' is the characteristic function of $Z(g)$ we have $g'g'' = 0 \in M'$ but $g' \notin M'$, hence $g'' \in M'$, a contradiction since $(g, g'') = C$.

The double correspondence thus established is in fact one-one. Starting with M, we have $f' \in M \cap C' \subset M'$ for any $f \in M$. Hence at least all of M, and therefore precisely M, returns on applying the second correspondence. Conversely, when we begin with M' we evidently have $M \cap C' \subset M'$ for the corresponding M. Since the expansion of $M \cap C'$ is unique, the second correspondence must return us to M'.

Finally to prove the correspondence a homeomorphism, it will suffice (since both spaces are compact) to prove that $M' \supset \cap M_i'$ implies $M \supset \cap M_i$. Suppose $f \in M_i$ for all i, then $f' \in M_i' \subset M'$, whence $f \in M$.

We remark that (for the case X discrete) the space of maximal ideals

[15] The field of rationals under the 6-adic topology is locally bounded but does not satisfy the latter condition.

in $C(X, A)$ is the same for any division ring A, and is in fact the Cech compactification βX.

Another interesting question is the following: if M' is a free maximal ideal in C', what is $C' - M'$? When A is the reals, complexes, or quaternions, or a finite field, it is known that $C' - M' = A$. Apparently it will require new methods to extend this result to more general cases.

UNIVERSITY OF CHICAGO.

BIBLIOGRAPHY.

1. W. Ambrose, "Structure theorems for a special class of Banach algebras," *Transactions of the American Mathematical Society*, vol. 57 (1945), pp. 364-386.

2. G. Birkhoff, "Lattice Theory," *Colloquium publications*, vol. 25 (New York, 1940).

3. I. S. Cohen, "On the structure and ideal theory of complete local rings," *Transactions of the American Mathematical Society*, vol. 59 (1946), pp. 54-106.

4. D. van Dantzig, "Zur topologische Algebra. II," *Compositio Mathematica*, vol. 2 (1935), pp .201-223.

5. M. Deuring, "Algebren," *Ergebnisse der Mathematik*, vol. 4 (1935).

6. S. Eilenberg and S. MacLane, "Group extensions and homology," *Annals of Mathematics*, vol. 43 (1942), pp. 757-831.

7. I. Gelfand, "Normierte Ringe," *Math. Sbornik*, vol. 9 (1941), pp. 1-23.

8. I. Gelfand and A. Kolmogoroff, "On rings of continuous functions on topological spaces," *Comptes Rendus* URSS (Doklady), vol. 22 (1939), pp. 11-15.

9. N. Jacobson and O. Taussky, "Locally compact rings," *Proceedings of the National Academy of Sciences*, vol. 21 (1935), pp. 106-108.

10. N. Jacobson, "Totally disconnected locally compact rings," *American Journal of Mathematics*, vol. 58 (1936), pp. 433-449.

11. N. Jacobson, "The radical and semi-simplicity for arbitrary rings," *American Journal of Mathematics*, vol. 67 (1945), pp. 300-320.

12. N. Jacobson, "A topology for the set of primitive ideals in an arbitrary ring," *Proceedings of the National Academy of Sciences*, vol. 31 (1945), pp. 333-338.

13. E. van Kampen, "Locally bicompact abelian groups and their character groups," *Annals of Mathematics*, vol. 36 (1935), pp. 448-463.

14. Y. Otobe, "On quasi-evaluations of compact rings," *Proceedings of the Imperial Academy of Tokyo*, vol. 20 (1944), pp. 278-282.

15. Y. Otobe, "On locally compact fields," *Japanese Journal of Mathematics*, vol. 19 (1945), pp. 189-202.

16. I. Shafarevich, "On the normalizability of topological fields," *Comptes Rendus URSS* (Doklady), vol. 40 (1943), pp. 133-135.

17. M. H. Stone, "Applications of the theory of Boolean rings to general topology," *Transactions of the American Mathematical Society*, vol. 41 (1937), pp. 375-481.

18. A. Weil, "L'intégration dans les Groupes Topologiques et ses Applications," *Actualités Scientifiques et Industrielles* 869, Paris, 1938.

UNIQUENESS OF SOLUTIONS OF ULTRAHYPERBOLIC PARTIAL DIFFERENTIAL EQUATIONS.*

By O. G. Owens.

Introduction. A solution in a bounded region V of the elliptic partial differential equation, $L[u] \equiv \sum_{i=1}^{n} u_{x_i x_i} = 0$, is unique if on the boundary of V the solution assumes prescribed values. The theorem is proved by transforming by means of Green's formula the integral

$$\int_V uL[u]dx.$$

A solution in a bounded region V of the normal hyperbolic partial differential equation, $L[u] \equiv u_{tt} - \sum_{i=1}^{n} u_{x_i x_i} = 0$, is unique if V is bounded by a characteristic cone of the equation and by a plane $t =$ constant, and if on the plane boundary of V the solution and its normal derivative assume prescribed values. The theorem is proved by transforming the integral

$$\int_V 2u_t L[u]dxdt. \quad [1]$$

The purpose of this paper is to establish an analogue of the above classical integral uniqueness procedures for the ultrahyperbolic partial differential equation,

$$(1) \qquad L[u] \equiv \sum_{i=1}^{n} u_{x_i x_i} - \sum_{k=1}^{m} u_{y_k y_k} = 0 \qquad (n \geq 2, m \geq 2),$$

The integral yielding the desired uniqueness conditions for solutions of (1) is:

$$(2) \qquad \int_V [2\sum_{l=1}^{n} x_l u_{x_l} + nu]L[u]dxdy.$$

The resulting uniqueness theorem is applicable to a limit problem which is

* Received April 2, 1945; Revised July 20, 1946.

not a Cauchy problem [2] but a mixed problem having an elliptic and hyperbolic nature.

Uniqueness Theorems. Three typical uniqueness theorems applying to solutions of (1) are:

THEOREM 1. *A solution of* (1) *existing in a hyper-sphere V is unique if the solution assumes prescribed values on the surface V^* of V and if its normal derivative assumes prescribed values on one of the two regions subdividing V^* and having as boundary the intersection of V^* and the characteristic cone of* (1) *with vertex at the center of V.*

THEOREM 2. *A solution of* (1) *existing in a hyper-parallelepiped with faces parallel to the co-ordinate planes is unique if the solution assumes prescribed values on the faces and if its normal derivative is prescribed on a single face.*

THEOREM 3. *A solution of* (1) *existing in a bounded region C is unique if the solution assumes prescribed values on the frontier of C and if C is the region defined by slicing, with planes: $x_l = 0$, $x_l + \sum_{k=1}^{m} b_k y_k = \text{const.}$ $(\sum_{k=1}^{m} b_k^2 \geqq 1)$, any hyper-cylinder having generators parallel to the x_l-axis.*

Theorem 1 will now be proved. If there are two solutions of (1) existing in V and assuming the same mixed initial data on V^*, then their difference is a solution assuming vanishing data on V^*. Hence it suffices to show that the latter solution vanishes identically in V. The proof of this fact is commenced by introducing the symbol

$$\delta_{il} = \left\{ \begin{array}{cc} 1 & i \neq l \\ -1 & i = l \end{array} \right\}$$

and forming the identity:

$$(3) \qquad \begin{aligned} 2x_l u_{x_l} L[u] \equiv & \sum_{i=1}^{n} \delta_{il} u_{x_i}^2 - \sum_{k=1}^{m} u_{y_k}^2 \\ & + \sum_{i=1}^{n} \left\{ \frac{\partial}{\partial x_i} (2x_l u_{x_l} u_{x_i}) - \frac{\partial}{\partial x_l} (x_l u_{x_i}^2) \right\} \\ & - \sum_{k=1}^{m} \left\{ \frac{\partial}{\partial y_k} (2x_l u_{x_l} u_{y_k}) - \frac{\partial}{\partial x_l} (x_l u_{y_k}^2) \right\}. \end{aligned}$$

Now denote an element of the surface V^* by dV^* and denote the direction cosines of the outwardly directed normal at dV^* by

$$\frac{\partial x_i}{\partial \nu}, \quad \frac{\partial y_k}{\partial \nu}.$$

Thus, integrating (3) over V,

$$(4) \qquad \int_V 2x_l u_{x_l} L[u]\,dxdy = \int_V \sum_{i=1}^{n} \delta_{i_l} u_{x_i}{}^2 dxdy - \int_V \sum_{k=1}^{m} u_{y_k}{}^2 dxdy$$
$$+ \int_{V^*} \sum_{i=1}^{n} \{2x_l \frac{\partial x_i}{\partial \nu} u_{x_l} u_{x_i} - x_l \frac{\partial x_l}{\partial \nu} u_{x_i}{}^2\} dV^*$$
$$- \int_{V^*} \sum_{k=1}^{m} \{2x_l \frac{\partial y_k}{\partial \nu} u_{x_l} u_{y_k} - x_l \frac{\partial x_l}{\partial \nu} u_{y_k}{}^2\} dV^*.$$

Furthermore, assuming the center of the sphere V to be the origin of coordinates and the radius of V to be one,

$$\frac{\partial x_i}{\partial \nu} = x_i, \quad \frac{\partial y_k}{\partial \nu} = y_k.$$

Hence, upon substituting these values in (4),

$$(5) \qquad \int_V 2x_l u_{x_l} L[u]\,dxdy = \int_V \sum_{i=1}^{n} \delta_{i_l} u_{x_i}{}^2 dxdy - \int_V \sum_{k=1}^{m} u_{y_k}{}^2 dxdy$$
$$- \int_{V^*} \sum_{i=1}^{n} [x_l u_{x_i} - x_i u_{x_l}]^2 dV^* + \int_{V^*} \sum_{i=1}^{n} x^2{}_i u_{x_l}{}^2 dV^*$$
$$+ \int_{V^*} \sum_{k=1}^{m} [x_l u_{y_k} - y_k u_{x_l}]^2 dV^* - \int_{V^*} \sum_{k=1}^{m} y_k{}^2 u_{x_l}{}^2 dV^*.$$

Therefore, summing (5) with respect to l,

$$(6) \qquad \int_V \sum_{l=1}^{n} 2x_l u_{x_l} L[u]\,dxdy = (n-1) \int_V \sum_{i=1}^{n} u_{x_i}{}^2 dxdy - n \int_V \sum_{k=1}^{m} u_{y_k}{}^2 dxdy$$
$$+ \sum_{l=1}^{n} \sum_{k=1}^{m} \int_{V^*} [x_l u_{y_k} - y_k u_{x_l}]^2 dV^* - \sum_{l=1}^{n} \sum_{i=1}^{n} \int_{V^*} [x_l u_{x_i} - x_i u_{x_l}]^2 dV^*$$
$$+ \int_{V^*} [\sum_{i=1}^{n} x_i{}^2 - \sum_{k=1}^{m} y_k{}^2][\sum_{l=1}^{n} u_{x_l}{}^2] dV^*.$$

In addition to relation (6) we have the following:

$$(7)\qquad \int_V uL[u]dxdy = \int_V \sum_{k=1}^{m} u_{y_k}{}^2 dxdy - \int_V \sum_{i=1}^{n} u_{x_i}{}^2 dxdy + \int_{V^*} u[\sum_{i=1}^{n} x_i u_{x_i} - \sum_{k=1}^{m} y_k u_{y_k}]dV^*.$$

Consequently, because of (6) and (7),

$$(8)\qquad \int_V [\sum_{l=1}^{n} 2x_l u_{x_l} + nu]L[u]dxdy = -\int_V \sum_{i=1}^{n} u_{x_i}{}^2 dxdy + \sum_{l=1}^{n}\sum_{k=1}^{m}\int_{V^*} [x_l u_{y_k} - y_k u_{x_l}]^2 dV^* - \sum_{l=1}^{n}\sum_{i=1}^{n}\int_{V^*} [x_l u_{x_i} - x_i u_{x_l}]^2 dV^* + n\int_{V^*} u[\sum_{i=1}^{n} x_i u_{x_i} - \sum_{k=1}^{m} y_k u_{y_k}]dV^* + \int_{V^*} [\sum_{i=1}^{n} x_i{}^2 - \sum_{k=1}^{m} y_k{}^2][\sum_{l=1}^{n} u_{x_l}{}^2]dV^*.$$

Thus, recalling the assumption that u is zero on V^* and noticing that $[x_l u_{y_k} - y_k u_{x_l}]$, $[x_l u_{x_i} - x_i u_{x_l}]$ are inner derivatives on V^*,

$$(9)\qquad \int_V [\sum_{l=1}^{n} 2x_l u_{x_l} + nu]L[u]dxdy = -\int_V \sum_{i=1}^{n} u_{x_i}{}^2 dxdy + \int_{V^*} [\sum_{i=1}^{n} x_i{}^2 - \sum_{k=1}^{m} y_k{}^2][\sum_{l=1}^{n} u_{x_l}{}^2]dV^*.$$

Now, as in the case of the normal hyperbolic partial differential equation, the prescribing of limit values for a solution of (1) is influenced by its characteristic cone:

$$(10)\qquad \sum_{i=1}^{n} x_i{}^2 - \sum_{k=1}^{m} y_k{}^2 = 0.$$

The effect in our case is that the normal derivative of the solution need be prescribed, as being zero, only on that portion of V^* for which

$$\sum_{i=1}^{n} x_i{}^2 - \sum_{k=1}^{m} y_k{}^2 \geqq 0.$$

Since this suffices for us to infer from (9) that

$$(11)\qquad \int_V \sum_{i=1}^{n} u_{x_i}{}^2 dxdy = 0.$$

Hence, since (11) implies that u does not depend on the variables x_i, u is a solution of the equation,

$$\sum_{k=1}^{m} u_{y_k y_k} = 0. \tag{12}$$

But (12) together with the fact that $u = 0$ on V^* assures the vanishing of u on V. Thus, Theorem 1 has been proved.

The proof of Theorem 2 will now be indicated. Firstly, it is no specialization to assume the parallelepiped V to be defined by the inequalities,

$$0 \leq x_i \leq a_i, \qquad 0 \leq y_k \leq b_k \qquad\qquad (a_i > 0, \quad b_k > 0),$$

and the normal derivative of the solution to be zero on the face $x_l = a_l$. Secondly, because of (5) and (7),

$$\int_V [2x_l u_{x_l} + u] L[u] dxdy = -\int_V u_{x_l}{}^2 dxdy,$$

granted that u vanishes on the faces of V and the normal derivative of u vanishes on the face $x_l = a_l$. Thus, since $L[u] \equiv 0$, the solution is independent of the variable x_l. Hence, as $u \equiv 0$ on the face $x_l = a_l$, $u \equiv 0$ on V.

The proof of Theorem 3 is carried through in an analogous manner.

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

UNIVERSITY OF NEVADA,
RENO, NEVADA.

REFERENCES.

[1] K. Friedrichs and H. Lewy, *Mathematische Annalen*, vol. 98 (1927-28), pp. 192-204.
[2] F. John, *Proceedings of the National Academy of Sciences*, vol. 29 (1943), pp. 98-104.

NOTE ON SOME STRONG LAWS OF LARGE NUMBERS.*

By KAI-LAI CHUNG.

Let $\{X_n\}$ denote a sequence of independent random variables with the distribution functions $\{V_n(x)\}$ and let $S_n = \sum_{\nu=1}^{n} X_\nu$. Let $\{a_\nu\}$ denote a positive increasing sequence tending to infinity.

Recently Feller [1] has proved some theorems which can be looked upon as being extensions of a theorem of Marcinkiewicz and Zygmund.[2] In the spirit of this extension, we shall give an analogous extension of another theorem of the latter authors.[3] together with some colloraries dealing with the strong law of large numbers.

For the sake of completeness we state the main result of Feller's paper in a version suitable for our purposes. It reads as follows:

THEOREM 1 (Feller). *Let* $V_n(x) = V(x)$. *Suppose that*

$$\sum_{\nu=i}^{\infty} 1/a_\nu^2 = O(i/a_i^2)$$

and that either

$$a_n/n \uparrow$$

or

$$a_n/n \downarrow \quad \text{and} \quad \int_{-\infty}^{\infty} x dV(x) = 0.$$

Let $\psi(x)$ *be a positive, even function such that* $\psi(x) \uparrow$ *for* $x > 0$ *and let*

$$\psi(a_\nu) = \nu;$$

then

$$\Pr\left(\sum_{\nu=1}^{\infty} X_\nu/a_\nu \text{ converges}\right) = \begin{cases} 0 \\ 1 \end{cases}$$

according as

$$\int_{-\infty}^{\infty} \psi(x) dV(x) \begin{cases} = \\ < \end{cases} \infty.$$

* Received August 7, 1946.

[1] Feller, W., "A limit theorem for random variables with infinite moments," *American Journal of Mathematics*, vol. 68 (1946), pp. 257-262.

[2] Marcinkiewicz, J. and Zygmund, A., "Sur les fonctions indépendantes," *Fundamenta Mathematicae*, vol. 29 (1937), pp. 60-90. See Theorem 5.

[3] *Loc. cit.*, see Theorem 9.

A close analogue of Theorem 1 for the more general case of unequal distributions is the following.

THEOREM 2. *Let $\psi(x)$ be a positive, even function non-decreasing for $x > 0$. Suppose that either*

(i) $\psi(x)/x \downarrow$

or

(ii) $\psi(x)/x \uparrow$ $\quad \psi(x)/x^2 \downarrow$ *and* $\int_{-\infty}^{\infty} x dV_\nu(x) = 0$ *for all* ν.

Let

$$\int_{-\infty}^{\infty} \psi(x) dV_\nu(x) = M_\nu^{(\psi)}. \tag{1}$$

If the series

$$\sum_{\nu=1}^{\infty} M_\nu^{(\psi)}/\psi(a_\nu) \tag{2}$$

is convergent, then

$$\Pr\Big(\sum_{\nu=1}^{\infty} X_\nu/a_\nu \text{ converges}\Big) = 1. \tag{3}$$

Conversely, if the series (2) *is divergent, then there exists a sequence of independent random variables* $\{X_\nu\}$ *having distribution functions* $\{V_\nu(x)\}$ *such that* (1) *is satisfied and*

$$\Pr\Big(\sum_{\nu=1}^{\infty} X_\nu/a_\nu \text{ diverges}\Big) = 1. \tag{4}$$

Proof. Let $E(X)$ denote the mathematical expectation of X. Define

$$Y_\nu = \begin{cases} X_\nu & \text{if } |X_\nu| < a_\nu, \\ 0 & \text{if } |X_\nu| \geqq a_\nu. \end{cases}$$

Consider the case (i) first. We have

$$\sum_{\nu=1}^{\infty} E(Y^2{}_\nu/a^2{}_\nu) = \sum_{\nu=1}^{\infty} \int_{|x|<a_\nu} (x^2/a^2{}_\nu) dV_\nu(x)$$

$$\leqq \sum_{\nu=1}^{\infty} \int_{|x|<a_\nu} (\psi(x)/\psi(a_\nu)) dV_\nu(x) \leqq \sum_{\nu=1}^{\infty} M_\nu^{(\psi)}/\psi(a_\nu)$$

since $\psi(a_\nu)/a_\nu{}^2 \leqq \psi(x)/x^2$ for $|x| < a_\nu$ by assumption (i).

Next we have

$$\sum_{\nu=1}^{\infty} |E(Y_\nu)|/a_\nu \leqq \sum_{\nu=1}^{\infty} \int_{|x|<a_\nu} (|x|/a_\nu) dV_\nu(x)$$

$$\leqq \sum_{\nu=1}^{\infty} \int_{|x|<a_\nu} (\psi(x)/\psi(a_\nu)) dV_\nu(x) \leqq \sum_{\nu=1}^{\infty} M_\nu^{(\psi)}/\psi(a_\nu).$$

Finally,

$$\sum_{\nu=1}^{\infty} \Pr(Y_\nu \neq X_\nu) = \sum_{\nu=1}^{\infty} \int_{|x|\geqq a_\nu} dV_\nu(x)$$

$$\leqq \sum_{\nu=1}^{\infty} \int_{|x|\geqq a_\nu} (\psi(x)/\psi(a_\nu)) dV_\nu(x) \leqq \sum_{\nu=1}^{\infty} M_\nu^{(\psi)}/\psi(a_\nu).$$

Hence (3) follows from Kolmogoroff's three-series theorem.[4]

In case (ii) the only modification is that we have

$$\sum_{\nu=1}^{\infty} |E(Y_\nu)|/a_\nu = \sum_{\nu=1}^{\infty} 1/a_\nu \left| \int_{|x|\geqq a_\nu} x dV_\nu(x) \right|$$

$$\leqq \sum_{\nu=1}^{\infty} \int_{|x|\geqq a_\nu} (|x|/a_\nu) dV_\nu(x) \leqq \sum_{\nu=1}^{\infty} 1/\psi(a_\nu) \int_{|x|\geqq a_\nu} \psi(x) dV_\nu(x)$$

since now $\psi(a_\nu)/a_\nu \leqq \psi(x)/x$ for $|x| \geqq a_\nu$ by assumption (ii).

Finally, if (2) is divergent, we proceed, following Kolmogoroff, by defining a sequence of independent $\{X_\nu\}$ as follows:

If $M_\nu^{(\psi)}/\psi(a_\nu) \leqq 1$,

$$X_\nu = \begin{cases} \pm a_\nu \text{ each with probability } M_\nu^{(\psi)}/2\psi(a_\nu); \\ 0 \quad \text{with probability} \quad 1 - M_\nu^{(\psi)}/\psi(a_\nu). \end{cases}$$

If $M_\nu^{(\psi)}/\psi(a_\nu) > 1$,

$$X_\nu = \pm \psi^{-1}(M_\nu^{(\psi)}) \text{ each with probability } 1/2,$$

where $\psi^{-1}(x)$ denotes the inverse function of ψ. In either case we have

$$\Pr(|X_\nu| \geqq a_\nu) \geqq \min(M_\nu^{(\psi)}/\psi(a_\nu), 1).$$

Hence by Borel's lemma

$$\Pr(\overline{\lim_{\nu\to\infty}} |X_\nu|/a_\nu \geqq 1) = 1,$$

and (4) follows.

---

[4] Kolmogoroff, A., "Über die Summen durch den Zufall bestimmten unabhängigen Grössen," *Mathematische Annalen*, vol. 99 (1928), pp. 309-319, and vol. 102 (1929), pp. 484-488.

By the well-known Kronecker theorem, we can deduce from (3) sufficient conditions for the strong law of large numbers. It is then seen that our criterion includes those of Kolmogoroff, Marcinkiewicz and Zygmund. Among the consequences, which seem to be new as far as we are aware, we point out the following cases.

COROLLARY 1. *Suppose* $\int_{-\infty}^{\infty} x dV_\nu(x) = 0$. *Let* $\phi(x)$ *be any positive, even function non-decreasing for* $x > 0$ *such that*

$$\sum_{\nu=1}^{\infty} 1/\nu\phi(\nu) < \infty.$$

In particular we may take $\phi(x) = |\log|x||^{1+\epsilon}$, $\epsilon > 0$. *Then if*

$$\int_{-\infty}^{\infty} |x| \phi(x) dV_\nu(x) \leqq C$$

where C *is independent of* ν, *we have*

$$\Pr(\lim_{n\to\infty} s_n/n = 0) = 1.$$

COROLLARY 2. *Suppose* $\int_{-\infty}^{\infty} x dV_\nu(x) = 0$. *If* $1 < p < 2$ *and for an* $\epsilon > 0$

$$\int_{-\infty}^{\infty} |x|^p |\log|x||^{1+\epsilon} dV_\nu(x) \leqq C$$

where C *is independent of* ν, *we have*

$$\Pr(\lim_{n\to\infty} s_n/n^{1/p} = 0) = 1.$$

These results may be compared with those for the case of equal distributions, where the conditions are weaker in that the logarithm factor in the above is omitted. Our results however are the best possible in the sense of the second part of Theorem 2.

PRINCETON UNIVERSITY,
PRINCETON, N. J.